Random Walks on Regular Languages and Algebraic Systems of Generating Functions

Steven P. Lalley

ABSTRACT. A *random walk on a regular language* is a Markov chain on the set of all finite words from a finite alphabet A whose transition probabilities obey the following rules: (1) Only the last two letters of a word may be modified in one jump, and at most one letter may be adjoined or deleted. (2) Probabilities of modification, deletion, and/or adjunction depend only on the last two letters of the current word. Special cases include (a) reflecting random walks on the nonnegative integers; (b) LIFO queues; (c) finite-range random walks on homogeneous trees; and (d) random walks on the modular group $PSL_2(\mathbb{Z})$. It is shown that the $n-$step transition probabilities of a random walk on a regular language must obey one of three different types of power laws. The analysis is based on the study of an algebraic system of generating functions related to the Green's function.

1. Introduction

The transition probabilities of random walks on the integer lattices, and more generally on countable abelian groups, satisfy *local limit theorems* that are intimately related to the *global* limit theorems, such as the CLT, that govern sums of i.i.d. random vectors. For random walks on countable nonabelian groups there are no such global limit theorems. Nevertheless, it is known that for random walks on certain countable nonabelian groups, including the finitely generated free groups [**20, 10, 14**] and the modular group $PSL(2,\mathbb{Z})$ [**22**], and for certain absolutely continuous random walks on semi-simple Lie groups [**3, 17**] the transition probabilities obey local limit theorems; and it is widely suspected that random walks on lattices of semi-simple Lie groups obey similar local laws.

This paper concerns random walks on another class of homogeneous structures, the *regular languages* of computer science [**12**]. Roughly, a *random walk on a regular language* is a Markov chain on the set of all finite words from a finite alphabet A whose transition probabilities obey the following rules: (1) Only the last two letters of a word may be modified in one jump, and at most one letter may be adjoined or deleted. (2) Probabilities of modification, deletion, and/or adjunction depend only on the last two letters of the current word. See section 2.1 for definitions, and section 3 for examples. It will be shown that, under suitable irreducibility and

aperiodicity hypotheses, the transition probabilities of such random walks must follow one of three different types of power laws: as $n \to \infty$,

$$(1.1) \qquad P(X_n = y \,|\, X_0 = x) \sim C_{xy} R^{-n}, \qquad \text{or}$$

$$(1.2) \qquad \sim C_{xy} R^{-n} n^{-1/2}, \qquad \text{or}$$

$$(1.3) \qquad \sim C_{xy} R^{-n} n^{-3/2}$$

for suitable constants $C_{xy} > 0$ and $R \geq 1$. (Here and throughout the paper the symbol $\sim$ indicates that the ratio of the two sides converges to one as $n \to \infty$.) Certain countable nonabelian groups, such as (i) the finitely generated free groups, (ii) free products of two-element groups, and (iii) the modular group, admit descriptions as regular languages, and for random walks on such groups local limit theorems will emerge as special cases. In addition, various other Markov chains which do not have representations as random walks on groups, such as LIFOs (last-in first out queues) with different job types, *do* have representations as random walks on regular languages, and so our results will imply that the transition probabilities of such chains must obey one of the three power laws.

The main tools to be used are Tauberian theorems for generating functions (in particular, Darboux' theorem [**2**]) and techniques for analyzing singularities of generating functions. The key generating functions, the *Green's functions* of the random walks, will be shown to satisfy recursive relations stemming from the recursive nature of a regular language. These recursive relations, in turn, will imply that a related finite system of generating functions satisfy a system of polynomial equations, and thus are algebraic functions. It will then follow that the singularities of the Green's functions are at a branch point on the Riemann surface determined by the algebraic system. Finally, it will be shown that the this branch point has ramification two; here the Perron-Frobenius theory of nonnegative matrices will play a crucial role.

The algebraic character of the Green's functions for random walk on a free group was first proved in [**1**]. This result was subsequently generalized to nearest-neighbor random walks on any finitely generated group containing free normal subgroups of finite index in [**23**], and it has been conjectured that these are the *only* groups for which the Green's functions of nearest-neighbor random walks are algebraic. The arguments of papers [**1**] and [**23**] give no information about the nature of the singularities of the Green's functions.

2. Random walk on a regular language

2.1. Definitions and Hypotheses. A *(nearest-neighbor) random walk on a regular language* (abbreviated *RWRL*) is a Markov chain X_n on the set $A^* := \cup_{n \geq 0} A^n$ of all finite words from a finite alphabet A whose transitions obey the following rules:

(a) Only the last two letters of the current word may be modified.
(b) At most one letter may be adjoined or deleted.
(c) Adjunction or deletion may only be done at the *end* of a word.
(d) Probabilities of modification, deletion, and/or adjunction depend only on the last two letters of the current word.

More precisely, a RWRL is a Markov chain whose transition probabilities satisfy:

$$(2.1) \qquad P(X_{n+1} = x_1 x_2 \ldots x_m a'b' \,|\, X_n = x_1 x_2 \ldots x_m ab) = p(a'b'|ab)$$
$$P(X_{n+1} = x_1 x_2 \ldots x_m a' \,|\, X_n = x_1 x_2 \ldots x_m ab) = p(a'|ab)$$
$$P(X_{n+1} = x_1 x_2 \ldots x_m a'b'c \,|\, X_n = x_1 x_2 \ldots x_m ab) = p(a'b'c|ab)$$
$$P(X_{n+1} = a'b' \,|\, X_n = a) = p(a'b'|a)$$
$$P(X_{n+1} = a' \,|\, X_n = a) = p(a'|a)$$
$$P(X_{n+1} = \emptyset \,|\, X_n = a) = p(\emptyset|a)$$
$$P(X_{n+1} = a \,|\, X_n = \emptyset) = p(a|\emptyset) \qquad \text{and}$$
$$P(X_{n+1} = \emptyset \,|\, X_n = \emptyset) = p(\emptyset|\emptyset).$$

Here $\emptyset$ denotes the empty word. We do *not* assume that the transition probabilities are all positive, nor do we assume that there exist positive-probability paths connecting any two words. We shall, however, restrict attention to those random walks satisfying the following aperiodicity and irreducibility hypotheses:

ASSUMPTION 2.1 (Aperiodicity). *For all words $x_1 x_2 \ldots x_m$,*

$$(2.2) \qquad P(X_{n+1} = x_1 x_2 \ldots x_m \,|\, X_n = x_1 x_2 \ldots x_m) > 0$$

Observe that, in view of the homogeneity restrictions implicit in equations (2.1), this is equivalent to assuming that all of the transition probabilities $p(ab \,|\, ab)$, $p(a \,|\, a)$, and $p(\emptyset \,|\, \emptyset)$ are positive.

ASSUMPTION 2.2 (Irreducibility-A). *Let $\mathcal{L} \subset A^*$ be the set of all words that may be reached from $\emptyset$ via positive-probability paths. Then $\forall w \in \mathcal{L}$ there is a positive-probability path from w to $\emptyset$.*

ASSUMPTION 2.3 (Irreducibility-B). *If*

$$w_1 = x_1 x_2 \ldots x_m a \in \mathcal{L} \qquad \text{and}$$
$$w_2 = x_1 x_2 \ldots x_m \in \mathcal{L}$$

then there exists a positive-probability path from w_1 to w_2 through words of length at least $m + 1$.

ASSUMPTION 2.4 (Irreducibility-C). *If*

$$w_1 = x_1 x_2 \ldots x_m a_1 a_2 \ldots a_k \in \mathcal{L} \qquad \text{and}$$
$$w_2 = b_1 b_2 \ldots b_{k'} y_1 y_2 \ldots y_{m'} \in \mathcal{L}$$

then there exists a word $w_3 \in \mathcal{L}$ of the form

$$w_3 = b_1 b_2 \ldots b_{k'} z_1 z_2 \ldots z_{m''} a_1 a_2 \ldots a_k.$$

Moreover, for each such triple w_1, w_2, w_3 there exists a positive-probability path from w_2 to w_3 through words of length $\geq k'$.

Henceforth, it is assumed that a RWRL satisfies the aperiodicity and irreducibility assumptions above. A number of interesting examples are discussed in section 3 below. It should be clear that our results may be extended to RWRLs satisfying somewhat weaker hypotheses. For instance, the hypothesis that changes of state should involve only the last two letters could be replaced by the weaker hypothesis that such changes involve only the last K letters, for some $K < \infty$: see the discussion in paragraph 3.3 below. The aperiodicity hypothesis (2.2) is mainly for ease of exposition — random walks on regular languages that do not satisfy

this hypothesis may have transition probabilities that oscillate periodically, with any integer period. See section 6 below for an example. Finally, the irreducibility assumptions 2.2-2.4 may be weakened; however, the assumptions hold in many interesting examples, including those discussed in section 3 below.

It is easily proved that the set $\mathcal{L}$ of words accessible from the root $\emptyset$ is a *regular language*, that is, a subset of A^* consisting of all words accepted by a finite-state automaton. See [**12**] for background on regular languages. Since our results and arguments will make no appeal to results from the theory of formal languages, we shall not discuss this point in any further detail; however, we note in passing that it is essentially the recursive structure of $\mathcal{L}$ that lies behind the system of polynomial equations (5.9) below, on which our analysis is based.

2.2. Local Limit Theorem. The main results of this paper are the following theorems, which concern the asymptotic behavior of the transition probabilities $P^x\{X_n = y\} := P(X_n = y \mid X_0 = x)$.

THEOREM 2.5. *Let X_n be a* null recurrent *RWRL. Then for each pair $x, y \in \mathcal{L}$ there exists a constant $0 < C_{xy} < \infty$ such that as $n \to \infty$,*

$$(2.4) \qquad P^x\{X_n = y\} \sim C_{xy}n^{-1/2}$$

THEOREM 2.6. *Let X_n be a* transient *RWRL. Then for each pair $x, y \in \mathcal{L}$ there exists a constant $0 < C_{xy} < \infty$ and $1 < R < \infty$ (R is the* spectral radius *of the random walk) such that one of the following holds as $n \to \infty$:*

$$(2.5) \qquad P^x\{X_n = y\} \sim C_{xy}R^{-n}; \ or$$

$$(2.6) \qquad P^x\{X_n = y\} \sim C_{xy}R^{-n}n^{-1/2}; \ or$$

$$(2.7) \qquad P^x\{X_n = y\} \sim C_{xy}R^{-n}n^{-3/2}.$$

The proofs of these theorems for the special case where x and y are words of length ≤ 1 will be carried out in sections 5, 7, and 8 below. The general case, where x, y are arbitrary, may be deduced from this special case by a "recoding" trick — see paragraph 3.3 below. It will be shown in section 6 that all three cases of Theorem 2.6 may occur. It is known (see, for instance, [**14**] for the proof) that for random walk on a nonamenable group the power laws (2.5) and (2.6) cannot occur. Thus, for the examples considered in sections 3.5 and 3.6 below, only (2.7) is possible.

It is somewhat mysterious that only the three possibilities (2.5)-(2.7) can occur. Although we have no purely probabilistic explanation, we believe that the latter two power laws correspond to different possible limit laws for the "pinned" random walks. Consider the renormalized wordlength process

$$(2.8) \qquad W_n(t) := |X_{\lfloor nt \rfloor}|/\sqrt{n},$$

where $|\cdot|$ denotes wordlength and $\lfloor\cdot\rfloor$ denotes greatest integer. We conjecture that the power laws (2.6) and (2.7) correspond, respectively, to the following possibilities for the processes $\{W_n(t)\}_{0 \leq t \leq 1}$:

$$(2.9) \qquad W_n(t) \xrightarrow{\mathcal{D}} \text{Brownian bridge, and}$$

$$(2.10) \qquad W_n(t) \xrightarrow{\mathcal{D}} \text{Brownian excursion.}$$

These are, of course, the only possibilities for the reflected random walks discussed in sections 3.1 and 6 below. Additional evidence for the conjecture is given in

a recent paper of BOUGEROL & JEULIN [**4**], where it is shown that the renormalized distance-to-the-identity processes for isotropic random walks on the hyperbolic plane must converge to the Brownian excursion.

2.3. Malyshev's Theorem. Random walks on regular language have previously been studied by MALYSHEV and co-authors in [**18**] and [**9**], in which they are called "random strings". Malyshev proves the following basic theorem:

THEOREM . The wordlength $L_n := |X_n|$ of a RWRL satisfies a SLLN:

$$(2.11) \qquad\qquad L_n/n \to \mu \qquad \text{almost surely as } n \to \infty,$$

and $\mu = 0$ if and only if X_n is recurrent. In all cases, there is "edge stabilization": for any $m \geq 1$, the joint distribution of the last m letters of X_n converges weakly as $n \to \infty$.

Theorems of this type, of course, give no information about the detailed asymptotic behavior of the transition probabilities.

3. Examples of random walks on regular languages

3.1. Nearest-neighbor reflecting random walks on $\mathbb{Z}_+$. The most basic example is the simple nearest-neighbor $p - q$ random walk with reflection at the the origin. This is the Markov chain on $\mathbb{Z}_+$ with transition probabilities

$$
\begin{aligned}
(3.1) \qquad P(X_{n+1} = x + 1 \,|\, X_n = x) &= p && \text{if } x \geq 1; \\
P(X_{n+1} = x - 1 \,|\, X_n = x) &= 1 - p && \text{if } x \geq 1; \\
P(X_{n+1} = 1 \,|\, X_n = 0) &= r; \\
P(X_{n+1} = 0 \,|\, X_n = 0) &= 1 - r
\end{aligned}
$$

Henceforth, we shall refer to this random walk as $RRW(p; r)$. This is clearly equivalent to a random walk on the regular language $\mathcal{L} = \{a^m\}_{m \geq 0}$ with a one-letter alphabet. Notice, though, that this random walk does not satisfy the aperiodicity hypothesis (2.2) and in fact, as we will show in section 6 below, the transition probabilities do not obey the local limit theorem (2.7), but rather exhibit a periodic oscillation.

3.2. *Non*-nearest-neighbor reflecting random walks on $\mathbb{Z}_+$. Consider a Markov chain on $\mathbb{Z}_+$ with transition probabilities

$$
\begin{aligned}
(3.2) \quad P(X_{n+1} = x + k \,|\, X_n = x) &= p_k && \text{for } |k| \leq K \text{ and } x \geq K; \\
P(X_{n+1} = x + k \,|\, X_n = x) &= p_{x,k} && \text{for } 0 \leq x < K \text{ and } -x \leq k \leq K.
\end{aligned}
$$

Such a process behaves as a (homogeneous) random walk with increments bounded by K outside a finite neighborhood $\{0, 1, \cdots, K-1\}$ of the origin. It is not difficult to see that a Markov chain governed by transition probabilities (3.2) is equivalent to a random walk on the regular language

$$\mathcal{L} = \{a^m b \,|\, m \in \mathbb{Z}_+ \text{ and } b \in \{0, 1, 2, \cdots, K - 1\}\}.$$

The word $a^m b$ corresponds to the integer $mK + b$. Since the increments in X_n are of magnitude $\leq K$, only the last two letters of the representing word $aa \cdots ab$ need be modified to account for any possible change of state. Local limit theorems for Markov chains with transition probabilities (3.2) were proved, by Wiener-Hopf methods, in [**15**].

3.3. The "recoding" trick. The hypothesis that transitions in a RWRL involve only the last two letters of the word representing the current state may seem at first glance to be quite restrictive. However, the trick used in the preceding paragraph to represent non-nearest neighbor reflecting random walks on $\mathbb{Z}_+$ as RWRLs may be adapted to various other Markov chains.

Suppose, for instance, that X_n is a Markov chain on A^* whose transition probabilities obey the restrictions (a)-(d) of section 2.1, *except* that in any given transition, the last K letters (instead of the last 2) may be modified, and the probabilities of adjunction, deletion, and modification depend on the last K letters (not just the last 2). Although the transition probabilities of the Markov chain X_n do not obey the rules (2.1) (unless $K = 2$), it is not difficult to construct a RWRL Y_n that "represents" X_n as follows: Replace the alphabet A by the larger alphabet $B := \cup_{m=1}^K A^m$ consisting of all words of length $\leq K$ from A. For each finite word x in the letters A, let $y = y(x)$ be the word in the letters B obtained from x by grouping letters of x in blocks of K, starting at the beginning of the word. For example, if $K = 3$:

$$x = a_1 a_2 a_3 a_4 \qquad \Longrightarrow$$
$$y(x) = (a_1 a_2 a_3)(a_4).$$

Observe that the mapping $x \mapsto y(x)$ is bijective. It is routine to verify that the transition probabilities of the Markov chain $Y_n = y(X_n)$ obey the rules (2.1).

Notice that the recoding trick just described also produces a RWRL when the input process X_n is already a RWRL. In this case, the recoding collapses words of length $\leq K$ to words of length 1. Therefore, to prove Theorems 2.5-2.6 for arbitrary x, y, it suffices to prove them for words x, y of length 0 or 1.

3.4. LIFO Queues with multiple job types. A LIFO (last-in first-out) *queue*, also called a *stack*, is a data structure used in many fundamental algorithms of computer science (see, for instance, [**5**], section 11). The queue is stored as a linked list. New items are adjoined at the top of the linked list ("push"); and when items are removed, they are removed, one at a time, from the top ("pop"). In certain algorithms (for instance, when the stack is used to implement a recursion), the items in the list represent *jobs* to be completed, one at a time. These jobs may be of different types. Assume that the set A of possible job types is finite.

Now suppose that inputs to the algorithm are random. Then during the execution of the algorithm the stack will evolve as a stochastic process whose state space is a subset of the set A^* of finite words over the alphabet A. In algorithms where the state of the computation depends only on the top job in the stack (as is often the case) and inputs to the algorithm are i.i.d., the sequence of successive states of the stack will constitute a random walk on a regular language.

3.5. Random walk on a finitely generated free group. The *free group* $\mathcal{G} = \mathcal{G}(B)$ generated by a finite set $B = \{b_1, b_2, \ldots, b_L\}$ consists of all *reduced* finite words in the letters $B^\pm = \{b_1^{\pm 1}, b_2^{\pm 1}, \ldots, b_L^{\pm 1}\}$. (A word is said to be *reduced* if no letter $b_i^{\pm 1}$ is adjacent to its inverse $b_i^{\mp 1}$.) Multiplication in $\mathcal{G}$ consists of concatenation followed by reduction; the empty word $\emptyset$ is the group identity. A *nearest-neighbor random walk* on $\mathcal{G}$ is a Markov chain X_n that evolves according to the rule

$$(3.3) \qquad\qquad\qquad X_{n+1} = X_n \xi_{n+1}$$

where $\xi_1, \xi_2, \ldots$ are i.i.d. random variables valued in $B^{\pm} \cup \{\emptyset\}$. Clearly, post-multiplication by an element of $B^{\pm}$ can modify only the last letter of a word in $\mathcal{G}$, and so the transition probabilities implied by (3.3) must satisfy the hypotheses (2.1). If each element $\alpha \in B^{\pm} \cup \{\emptyset\}$ is assigned positive probability by the distribution of ξ_n, then the aperiodicity and irreducibility hypotheses of section 2.1 will be satisfied.

More generally, consider a *non*-nearest-neighbor random walk on $\mathcal{G}$ whose increments are bounded. The law of evolution is again given by (3.3), but now it is assumed that the jumps ξ_n are i.i.d. random variables valued in the set of all reduced words of length $\leq K$. These random walks may also be represented as random walks on a regular language, using the recoding trick of paragraph 3.3 above. The appropriate regular language $\mathcal{L}$ consists of all finite words $a_1 a_2 \ldots a_m b$, with $m \geq 0$, such that (i) each $a_i \in \mathcal{G}$ is a word of length K (in the generators of the group $\mathcal{G}$); (ii) $b \in \mathcal{G}$ is a word of length $0 \leq |b| \leq K - 1$; and (iii) no cancellations between successive a_i, a_{i+1} (or between a_m, b) are possible. It is easily verified that if the distribution of ξ_1 attaches positive probability to each word of length $\leq K$ then the aperiodicity and irreducibility hypotheses of section 2.1 are satisfied.

Finally, observe that if the transition probabilities are modified at finitely many elements of $\mathcal{G}$, then the resulting Markov chain, a *perturbed random walk*, although no longer a "random walk" (in the usual sense of the term), is nevertheless still a random walk on a regular language. If the modifications are such that positive-probability transitions remain positive-probability transitions, then the aperiodicity and irreducibility hypotheses will still hold. However, whereas the power laws (2.5) and (2.6) cannot occur for processes of the form (3.3), they *can* occur for perturbed random walks.

3.6. Random walk on the modular group $PSL(2, \mathbb{Z})$. The *modular group* is the quotient group $PSL(2, \mathbb{Z}) = SL(2, \mathbb{Z})/\{\pm I\}$ where $SL(2, \mathbb{Z})$ is the (multiplicative) group of 2×2 matrices with integer entries and determinant 1, and I is the 2×2 identity matrix. (For information about the modular group and its connections with the theory of elliptic modular forms, and also for specific facts used in the subsequent discussion, see [**16**], especially Chapter XI.) A *(right) random walk* on $PSL(2, \mathbb{Z})$ is a sequence

$$(3.4) \qquad X_n = \xi_1 \xi_2 \ldots \xi_n$$

where $\xi_1, \xi_2, \ldots$ are i.i.d. random variables valued in $PSL(2, \mathbb{Z})$. We shall assume that (a) the distribution of ξ_1 has finite support; (b) $P\{\xi_1 = I\} > 0$; and (c) that for every element $A \in PSL(2, \mathbb{Z})$, the random walk $\{X_n\}_{n \geq 0}$ visits A with positive probability. We shall show that any such random walk has a description as a RWRL. The assumptions (b)-(c) above guarantee that the irreducibility and aperiodicity assumptions 2.1, 2.2, and 2.3 will be satisfied.

The modular group is finitely generated, with generators

$$(3.5) \qquad U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad U^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \qquad \text{and } T = T^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

(More accurately, the group is generated by the *equivalence classes* $\langle T \rangle := \{\pm T\}$ and $\langle U \rangle := \{\pm U\}$. Henceforth, we shall not distinguish between matrices and their equivalence classes in $PSL(2, \mathbb{Z})$.) These generators satisfy the relations

$$(3.6) \qquad T^2 = (TU)^3 = I.$$

Thus, every element of $PSL(2, \mathbb{Z})$ may be written as a finite word in the letters T, U, U^{-1} in which no two T's appear consecutively, and no U is adjacent to a U^{-1}. It follows that the modular group is also generated by T, W, W^2, where $W = TU$. The relations (3.6) translate as $T^2 = W^3 = I$. Every element $M \in PSL(2, \mathbb{Z})$ has a representation as a finite word

$$(3.7) \qquad M = T^a W^{n_1} T W^{n_2} T W^{n_3} \cdots T W^{n_k} T^b$$

where $a, b = 0$ or 1 and each n_i is either 1 or 2.

The modular group contains free normal subgroups of finite index, and so it has a representation as a regular language. Following is a sketch of how such a representation may be obtained (see [16], section XI.3E for details). Let Γ' be the *commutator subgroup* of $PSL(2, \mathbb{Z})$, that is, the subgroup consisting of all finite products of *commutators*. (The *commutators* of any group are the elements of the form $aba^{-1}b^{-1}$.) The commutator subgroup of any finite or countable group is a normal subgroup, and so Γ' is a normal subgroup of $PSL(2, \mathbb{Z})$. The group Γ' is free (see [16], Theorem XI.3E) , with generators

$$(3.8) \qquad \begin{aligned} X &= TWTW^2, & X^{-1} &= WTW^2T, \\ Y &= TW^2TW, & \text{and} \qquad Y^{-1} &= W^2TWT. \end{aligned}$$

The index of Γ' in $PSL(2, \mathbb{Z})$ (that is, the number of distinct cosets of Γ') is 6. Thus, since Γ' is a normal subgroup, one may choose 6 distinct elements $A_1, A_2, \ldots, A_6$ of $PSL(2, \mathbb{Z})$, one from each coset, such that every element of $PSL(2, \mathbb{Z})$ has a representation

$$(3.9) \qquad M = W A_i,$$

where W is a reduced word in the (free) generators X, X^{-1}, Y, Y^{-1}. Observe that there may be some cancellation of letters at the end of W; but since there are only 6 possibilities for A_i, there are only finitely many possible types of cancellation, involving at most a bounded number of letters at the end of W. Thus, there is a finite subset $\mathcal{B} \subset PSL(2, \mathbb{Z})$ such that every element of $PSL(2, \mathbb{Z})$ has a representation as a *reduced* word

$$(3.10) \qquad M = W' B_j,$$

where W' is a reduced word in X, X^{-1}, Y, Y^{-1} and $B_j \in \mathcal{B}$. This exhibits $PSL(2, \mathbb{Z})$ as a regular language $\mathcal{L}$, with alphabet $\mathcal{A} = \{X, X^{-1}, Y, Y^{-1}\} \cup \mathcal{B}$.

Now consider a right random walk (3.4) whose jumps ξ_i satisfy the assumptions (a), (b), and (c) above. By (a), the distribution of ξ_i has finite support; consequently, there are only finitely many different ways that multiplication by ξ_{n+1} can modify the end of the word X_n. In particular, there is a finite integer K such that all positive-probability transitions $X_n \to X_{n+1}$ involve only the last K letters of X_n, and such that the transition probabilities depend only on the last K letters of the current state X_n. By replacing the regular language $\mathcal{L}$ by a regular language $\mathcal{L}'$ whose alphabet consists of the words of length $\leq K$ in $\mathcal{L}$, one obtains a representation of the random walk X_n as a random walk on a regular language whose transition probabilities obey the restrictions (2.1) of section 2.1.

A local limit theorem for random walk on the modular group was proved by Woess [22], using different and somewhat more difficult arguments than those employed below. Later [23], Woess also proved that, for any nearest-neighbor random walk on a discrete group containing a free normal subgroup of finite index, the

Green's function is algebraic; however, his arguments do not allow one to deduce the type of the leading singularity.

4. Darboux' Theorem

The asymptotic behavior of the transition probabilities $P^x(X_n = y)$ is governed by the behavior of the generating function $\sum_n z^n P^x(X_n = y)$ near its smallest positive singularity. To extract the information about the transition probabilities from the behavior of this generating function, we shall appeal to a Tauberian theorem of DARBOUX [2].

DARBOUX' THEOREM . Let $G(z) = \sum_{n=0}^{\infty} a_n z^n$ be a power series with radius of convergence R. Suppose that G has no singularities in the closed disk $|z| \le R$ except $z = R$, and that in a neighborhood of $z = R$,

$$(4.1) \qquad G(z) = A(z)(1 - z/R)^\alpha + B(z)$$

where $A(z)$ and $B(z)$ are analytic near $z = R$, $A(R) \ne 0$, and $\alpha \in \mathbb{R} \setminus \{0, -1, -2, \ldots\}$. Then as $n \to \infty$,

$$(4.2) \qquad a_n \sim \frac{A(R)}{\Gamma(-\alpha) n^{1+\alpha} R^n}$$

More generally, if G has finitely many singularities $\zeta_1, \zeta_2, \ldots, \zeta_m$ on the circle $|z| = R$ of convergence, near each of which it has a representation

$$(4.3) \qquad G(z) = A_i(z)(1 - z/\zeta_i)^{\alpha_i} + B_i(z)$$

with $A_i(z)$ and $B_i(z)$ analytic near $z = \zeta_i$ and $\alpha_i \in \mathbb{R} \setminus \{0, -1, -2, \ldots\}$, then as $n \to \infty$,

$$(4.4) \qquad a_n \sim \sum_{i=1}^{m} \frac{A_i(\zeta_i)}{\Gamma(-\alpha_i) n^{1+\alpha_i} \zeta_i^n},$$

provided the right side is not 0.

The hypotheses of Darboux' theorem are tailored to power series that define *algebraic functions*. A complex function $G(z)$ defined for $z \in \Omega$, where Ω is a nonempty open subset of $\mathbb{C}$, is said to be *algebraic* (more precisely, a *branch* of an algebraic function) if there is an irreducible polynomial $P(z, w)$ in two variables such that

$$(4.5) \qquad P(z, G(z)) = 0 \qquad \forall\, z \in \Omega$$

Observe that if the polynomial relation (4.5) holds in Ω, then it holds for any analytic continuation of G to a domain containing Ω, by the Law of Permanence of functional equations. See [11] for a thorough discussion of algebraic functions, including a proof of the following fundamental result:

PROPOSITION . If $G(z)$ is an algebraic function, then it has only finitely many singularities in the (extended) complex plane, and near each singularity has the form (4.3), with rational exponent α.

If $G(z) = \sum_{n=0}^{\infty} a_n z^n$ is an algebraic function with a singularity at $z = R$ and no other singularities in $|z| \le R$, then the conclusion (4.2) holds *provided* the function $A(z)$ in the representation (4.1) is not zero at $z = R$. Rather than verify

this directly, we shall find it more convenient to work with leading term of the *Puiseux series* of G at $z = R$. The Puiseux series (see [**11**]) is

$$(4.6) \qquad\qquad G(z) = G(R) + \sum_{n=1}^{\infty} b_n (R - z)^{n/k},$$

where $k \geq 1$ is an integer; the series $\sum_{n \geq 1} b_n z^n$ has positive radius of convergence. Using the Darboux theorem, one may easily check that the asymptotic behavior of the coefficients a_n is determined by the exponent of the first (nonconstant) *nonzero* term in the Puiseux expansion. Thus, we have the following simplification in the hypotheses of Darboux' theorem:

COLLARY 4.1. *Let* $G(z) = \sum_{n=0}^{\infty} a_n z^n$ *be a power series with nonnegative coefficients* a_n *and radius of convergence* R. *Suppose that the function* $G(z)$ *has no singularities in the closed disk* $|z| \leq R$ *except* $z = R$, *and that as* $z \to R-$ *(through real arguments less than* R*) either*

$$(4.7) \qquad\qquad\qquad G(z) \sim C(1 - z/R)^{\alpha} \qquad or$$

$$(4.8) \qquad\qquad G(R) - G(z) \sim C(1 - z/R)^{\alpha},$$

with $C > 0$. *If* $G(z)$ *is an algebraic function, then relation (4.2) holds with* $A(R)$ *replaced by* C.

5. Green's functions and related generating functions

5.1. The Green's Functions. Assume now that X_n is a random walk on a regular language $\mathcal{L}$ satisfying the aperiodicity and irreducibility hypotheses of section 2.1. For $x, y \in \mathcal{L}$, define the *Green's function* $G_{xy}(z)$ as follows:

$$(5.1) \qquad G_{xy}(z) := \sum_{n=0}^{\infty} P^x\{X_n = y\} z^n = E^x \left(\sum_{n=0}^{\infty} z^n \mathbf{1}\{X_n = y\} \right).$$

Since the coefficients in the power series are all probabilities, the series converge absolutely and uniformly in the closed unit disc $|z| \leq 1$. Thus, the power series of each of the Green's functions has radius of convergence at least 1. In section 7.2 below, we shall prove that all of these power series have the same radius R of convergence. Since their coefficients are nonnegative, and, by Assumption 2.2, at least some are strictly *positive*, the following must hold:

(a) $G_{xy}(z)$ and its derivatives are positive for $0 < z < R$.
(b) $G_{xy}(z)$ is strictly increasing and strictly convex in the interval $[0, R)$.
(c) $\lim_{z \to R-} G_{xy}(z) = G_{xy}(R) \in (0, \infty]$.
(d) $G_{xy}(z)$ has a singularity at $z = R$.

Property (d) is PRINGSHEIM's theorem (see [**11**]).

The Green's functions $G_{xy}(z)$ are intimately related to, and in fact completely determined by, a finite system of generating functions $H_{ab,c}(z)$ indexed by words ab and c of lengths 1 and 2, respectively. These are defined as follows:

$$(5.2) \qquad H_{ab,c}(z) := E^{ab} z^T \mathbf{1}\{X_T = c\} = \sum_{n=1}^{\infty} z^n P^{ab}\{T = n \text{ and } X_n = c\}$$

where

$$(5.3) \qquad\qquad\qquad T := \min\{n \ : \ |X_n| < |X_0|\}.$$

Observe that some of the generating functions $H_{ab,c}(z)$ may be identically zero, as there may be no positive-probability paths from ab to c through words of length ≥ 2. However, Assumptions 2.2-2.4 imply that at least some of the functions $H_{ab,c}(z)$ are *not* identically zero. We shall refer to these as the *nondegenerate* first-passage generating functions. In section 7.1 below we will show that the nondegenerate first-passage generating functions $H_{ab,c}(z)$ have a common radius ρ of convergence, and in section 7.3 that $\rho \geq R$. Because the coefficients in the power series are nonnegative, and, for the nondegenerate generating functions, not all zero, the nondegenerate generating functions $H_{ab,c}(z)$ satisfy properties (a)-(d) above, with R replaced by ρ.

The relationship between the Green's functions and the first-passage generating functions $H_{ab,c}(z)$ is a simple consequence of the Markov property. Observe that the second equality in (5.1) shows that the Green's function $G_{xy}(z)$ may be interpreted as the *expected discounted number of visits to y* when the initial state is x. Let x and y be words of lengths $|x|, |y| \leq 1$. If $X_0 = x$, then X_1 must be a word of length $0, 1$, or 2; if X_1 is a word of length 2, then there can be no more visits to y until after the Markov chain X_n returns from the set of words of length ≥ 2. Hence, by the Markov property, if $|x| \leq 1$ and $|y| \leq 1$, then

$$(5.4)\quad G_{xy}(z) = \delta_x(y) + z \sum_{w:|w|\leq 1} p(w|x)G_{wy}(z) + z \sum_{ab}\sum_{c} p(ab|x)H_{ab,c}(z)G_{cy}(z),$$

where $\sum_{ab}$ is over all two-letter words, and $\sum_c$ is over all one-letter words.

The equations (5.4) may be written compactly in matrix form as follows. Define $\mathbf{G}(z)$ to be the matrix-valued function of z with entries $G_{xy}(z)$, indexed by words x, y of lengths ≤ 1, and define $\mathbf{P}$ to be the matrix of transition probabilities $p(y|x)$, where again x and y are words of length ≤ 1. Define $\mathbf{K}(z)$ to be the matrix-valued function of z with entries $K_{xy}(z)$, indexed by words x, y of length ≤ 1, defined by

$$(5.5)\qquad\qquad K_{xy}(z) = \sum_{ab} p(ab|x)H_{ab,y}(z) \qquad \text{if } |y| = 1;$$
$$= 0 \qquad\qquad\qquad \text{if } |y| = 0.$$

Then equations (5.4) are equivalent to the matrix equations

$$(5.6)\qquad\qquad \mathbf{G}(z) = \mathbf{I} + z\mathbf{P}\mathbf{G}(z) + z\mathbf{K}(z)\mathbf{G}(z) \qquad \Longleftrightarrow$$
$$\mathbf{I} = (\mathbf{I} - z\mathbf{P} - z\mathbf{K}(z))\,\mathbf{G}(z)$$

where $\mathbf{I}$ is the identity matrix. Notice that for all z of sufficiently small modulus $|z|$, the matrix $(z\mathbf{P} + z\mathbf{K}(z))$ has norm less than one, and so the matrix $(\mathbf{I} - z\mathbf{P} - z\mathbf{K}(z))$ is invertible; thus,

$$(5.7)\qquad\qquad \boxed{\mathbf{G}(z) = (\mathbf{I} - z\mathbf{P} - z\mathbf{K}(z))^{-1}}$$

for all z in a disc centered at $z = 0$.

5.2. Algebraic relations among the generating functions. Equations (5.6) above show that the generating functions G_{xy} and $H_{w,v}$ are algebraically interrelated. In this section, we will show that the auxiliary generating functions $H_{w,v}$ are themselves algebraically interrelated via a system of quadratic equations that define an algebraic curve. This will imply that these functions are algebraic. This, together with the equations (5.6) above, will imply that the Green's functions $G_{xy}(z)$ are algebraic functions of z.

Recall (see equations 2.1)) that, for a random walk on a regular language, only the last two letters of the current word play a role in the transition to the next word. Consequently, for any initial state $\xi = x_1 x_2 \ldots x_m ab \in \mathcal{L}$, the letters $x_1 x_2 \ldots x_m$ are ignored up to time T. It follows that for any word $\xi = x_1 x_2 \ldots x_m ab \in \mathcal{L}$, and any $c \in A$,

$$(5.8) \qquad H_{ab,c}(z) = E^\xi z^T \mathbf{1}\{X_T = x_1 x_2 \ldots x_m c\}.$$

The homogeneity relations (5.8) lead directly to a system of quadratic equations relating the functions $H_{ab,c}$. These equations derive from the Markov property: Starting from any word w, the random walk must make an initial jump to another state. If this first jump is to a word whose length is less than that of the initial word w, then $T = 1$; otherwise, the initial jump must be either to a word w' of the same length, or to a word w'' with one additional letter. In the latter case, T is the first time after the initial jump that *two* letters are removed; in the former, T is the first time that a single letter is removed. Summing over all possible initial jumps and applying the identity (5.8) yields the following system of equations:

$$(5.9) \qquad H_{ab,c}(z) = zp(c|ab) + zp(ab|ab)H_{ab,c}(z) + z \sum_{d,e \in A} p(de|ab)H_{de,c}(z)$$

$$+z \sum_{d,e,f \in A} p(def|ab) \sum_{g \in A} H_{ef,g}(z)H_{dg,c}(z),$$

We shall abbreviate this system of equations as

$$(5.10) \qquad \boxed{\mathbf{H}(z) = z\mathbf{Q}(\mathbf{H}(z))}$$

where $\mathbf{H}(z)$ is the vector of *nondegenerate* first-passage generating functions $H_{ab,c}(z)$ (that is, those that are *not* identically zero), and $\mathbf{Q}$ is the vector of quadratic polynomials $Q_{ab,c}$ in the variables $H_{ab,c}$ occurring on the right sides of the equations (5.9) above. Observe that $\mathbf{Q}$ is quadratic (that is, each of its components is quadratic) and does not depend on the variable z. Moreover, the coefficients in the polynomials $Q_{ab,c}$ are *positive*. This will be of crucial importance later.

5.3. Algebraic curves and the algebraic character of the Green's functions.
An *affine (algebraic) variety* $V \subset \mathbb{C}^n$ is defined to be the set of all zeros $(a_1, a_2, \ldots, a_n$ of a finite set of polynomials in n variables over the field $\mathbb{C}$ of complex numbers. (See, for instance [**6**], Ch. I, for the definition. Some authors use a different terminology, calling what we have defined to be an affine variety an *algebraic set*, and reserving the term *variety* for an *irreducible* algebraic set.) The system of equations (5.10) determines an affine variety V_H, since it consists of polynomial (in fact, quadratic) equations in the variables z and $H_{ab,c}$. Similarly, the system consisting of equations (5.10) and (5.6) determines a variety V_{GH}. (Observe that V_H is obtained by projecting the variety V_{GH} onto the z and $H_{ab,c}$ coordinates.) Notice that the point $(0, \mathbf{H}(0))$ lies on the variety V_H, and that the point $(0, \mathbf{H}(0), \mathbf{G}(0))$ lies on V_{GH}.

A basic theorem of algebraic geometry (see [**6**], section 4.6) states that any affine variety V has a decomposition $V = \cup_{i=1}^r V_i$ into irreducible affine varieties, where a variety W is termed *irreducible* if the only decompositions $W = U_1 \cup U_2$ of W into two varieties are such that $W = U_1$ or $W = U_2$.

PROPOSITION 5.1. *There exists an open neighborhood $\mathcal{N}$ of the point*

$$(0, \mathbf{H}(0), \mathbf{G}(0))$$

such that $\mathcal{N} \cap V_{GH}$ is contained in an irreducible algebraic curve *(a variety of dimension 1).*

PROOF. See [**6**], Chapter 9, for the definition of the (algebraic) dimension of an affine variety. For our purposes, it will only be important that the algebraic dimension coincides with the *transcendence degree* (see the proof of Corollary 5.2 below) of the field of rational functions over $\mathbb{C}$. The system of equations consisting of (5.10) and (5.6) has the form

$$(5.11) \qquad\qquad \mathbf{F}(z, \mathbf{w}) = \mathbf{0},$$

where $\mathbf{w} = (w_1, w_2, \ldots, w_m)$ and $\mathbf{F}$ is an $m-$ vector of polynomials in the variables z and $w_1, w_2, \ldots, w_m$. If $(z_*, \mathbf{w}_*)$ is a solution of the equation (5.11), then it is said to be a *nonsingular* point if the Jacobian matrix

$$(5.12) \qquad\qquad J := \frac{\partial \mathbf{F}}{\partial \mathbf{w}}$$

is nonsingular at $(z, \mathbf{w}) = (z_*, \mathbf{w}_*)$. If $(z_*, \mathbf{w}_*)$ is a nonsingular point of the variety $V_{\mathbf{F}}$ defined by the system (5.11), then, by Theorem 9.6.9 of [**6**], it lies on a unique irreducible component of the variety $V_{\mathbf{F}}$ of dimension 1. For the system consisting of equations (5.10) and (5.6), the function $\mathbf{F}$ may be written $\mathbf{F} = (\mathbf{F_H}, \mathbf{F_G})$, where

$$\mathbf{F_H}(z, \mathbf{H}) = \mathbf{H} - z\mathbf{Q}(\mathbf{H}) \qquad \text{and}$$
$$\mathbf{F_G}(z, \mathbf{H}, \mathbf{G}) = (\mathbf{I} - z\mathbf{P} - z\mathbf{K}(z))\,\mathbf{G}(z).$$

The Jacobian matrix at $(0, \mathbf{H}(0), \mathbf{G}(0))$ has the form

$$J = \begin{pmatrix} I & \mathbf{0} \\ * & I \end{pmatrix},$$

where I denotes the identity matrix and $\mathbf{0}$ the matrix all of whose entries are 0. This is certainly nonsingular, so the point $(0, \mathbf{H}(0), \mathbf{G}(0))$ is a nonsingular point of V_{GH}, and it lies on a unique irreducible component of dimension 1. $\qquad\square$

Remark. Observe that the condition (5.12) is also the hypothesis of the (complex) *Implicit Function Theorem* (see, for example, [**13**], Chapter 3). This theorem states that, if the Jacobian matrix J is nonsingular at $(z_*, \mathbf{w}_*)$ then, for each z in a neighborhood of z_*, there is a *unique* solution $\mathbf{w} = \mathbf{w}(z)$ of (5.11), and the mapping $z \mapsto \mathbf{w}(z)$ is analytic. In particular, the locus of solutions to (5.11) near $(z_*, \mathbf{w}_*)$ is analytically parametrized by a single parameter z; in other words, it is a complex manifold of (topological) dimension one. The theorem 9.6.9 of [**6**] cited in the proof above states, in essence, that the *topological* and *algebraic* dimensions of the variety coincide at nonsingular points. $\qquad\square$

COLLARY 5.2. *Each of the functions G_{xy} and $H_{ab,c}$ occurring as entries of the Green's function matrix $\mathbf{G}(z)$ and the first-passage generating function vector $\mathbf{H}(z)$ is algebraic.*

PROOF. Let V be the irreducible component the variety V_{GH} containing the point $(0, \mathbf{H}(0), \mathbf{G}(0))$. Each of the coordinate variables z, $H_{ab,c}$, and G_{xy} is a member of the field $\mathbb{C}[V]$ of *rational functions* on the variety V (see [**6**]), section 5.5). By Theorem 9.5.2 of [**6**], the *transcendence degree* of the field $\mathbb{C}[V]$ (that is, the maximal number of algebraically independent elements) coincides with the algebraic dimension of V, which, by Proposition 5.1, is 1. Thus, there is a polynomial relation between any two elements of $\mathbb{C}[V]$, in particular, between z and $H_{ab,c}$, and between z and G_{xy}. □

In principle, a polynomial $\mathcal{P}(z, w)$ relating z and any one of the Green's functions can be obtained by an "elimination" algorithm, starting from the system (5.9) and (5.6). See [**6**] for an elementary discussion of elimination theory. Unfortunately, in all but the very simplest cases, the system (5.9) is either too large for the computation to be carried out, or leads to a polynomial $\mathcal{P}$ of too high a degree to be solved by quadratures for $G_{xy}(z)$ in terms of z. There is, however, one interesting case where the system may be solved explicitly — this is discussed in the next section.

6. A Special Case: Reflecting Random Walk on $\mathbb{Z}_+$

This special case is instructive, in that explicit forulas for the generating functions may be obtained, and all of the possibilities enumerated in Theorems 2.5 and 2.6 are exhibited. However, the results of this section will play no role in the subsequent sections.

6.1. Explicit expressions for the Green's functions. The Green's functions of the random walks $RRW(p; r)$ defined in section 3.1 may be written in closed form as algebraic functions of their arguments. Let X_n be the Markov chain on the set of nonnegative integers with transition probabilities given by (3.1). Define the generating functions $G(z), F(z)$ by

$$(6.1) \qquad G(z) = \sum_{n=0}^{\infty} z^n P^0\{X_n = 0\} \qquad \text{and}$$

$$(6.2) \qquad F(z) = E^1 z^T \qquad \text{where}$$

$$(6.3) \qquad T = \min\{n : X_n = 0\}$$

By (5.4) and the Markov property, (see [**7**], section XIV.4, for similar calculations),

$$(6.4) \qquad G(z) = (1 - z(1 - r) - zrF(z))^{-1} \qquad \text{and}$$

$$(6.5) \qquad F(z) = (1 - p)z + pzF(z)^2.$$

Notice that the algebraic system (5.10) reduces to the single quadratic equation (6.5). This equation has two solutions; since the function $F(z)$ defined by (5.2) is nondecreasing in z for $0 \le z \le 1$, the correct branch is

$$(6.6) \qquad F(z) = \frac{1 - \sqrt{1 - 4p(1-p)z^2}}{2pz}.$$

The smallest positive singularity is at $z = \rho$ where

$$(6.7) \qquad \rho = \rho(p) = 1/(2\sqrt{p(1-p)}).$$

This is a branch point of order two. Observe that $-\rho$ is also a branch point of order two, but that there are no other singularities on the circle $|z| = \rho$. Notice also that the function $H(\zeta) := \sqrt{\zeta}F(\sqrt{\zeta})$ has only one singularity on its circle of convergence, at $\zeta = (4p(1-p))^{-1}$, and so by Darboux' theorem, as $n \to \infty$ through the *odd* integers,

$$(6.8) \qquad P^1\{T = n\} \sim A(p;r)\rho^{-n}n^{-3/2}$$

for a suitable value $A(p;r) \in (0, \infty)$.

6.2. The Case $p = 1/2$. The reflecting random walk $RRW(1/2; r)$ is null recurrent for every value of $r \in (0,1)$. The equation (6.4) for $G(z)$ simplifies to

$$(6.9) \qquad G(z) = \left(1 - z(1-r) - r + r\sqrt{1-z^2}\right)^{-1},$$

which shows that G has singularities at $z = \pm 1$. These singularities are of different types, however, as $G(z) \to 0$ as $z \to 1$, while $G(z) \to 2 - r \neq 0$ as $z \to -1$. It is routine to check that

$$(6.10) \qquad G(z) = A_+(z)(1-z)^{-1/2} + B_+(z) \qquad \text{near } z = 1 \qquad \text{and}$$

$$(6.11) \qquad G(z) = A_-(z)(1-z)^{+1/2} + B_-(z) \qquad \text{near } z = -1$$

for analytic functions $A_\pm, B_\pm$. Thus, by Darboux' theorem, as $n \to \infty$ the return probabilities satisfy

$$(6.12) \qquad P^0\{X_n = 0\} \sim \frac{C}{\sqrt{n}}$$

for a suitable constant $C = C(1/2; r) > 0$.

6.3. The Case $p > 1/2$. When $p > 1/2$ the reflecting random walk $RRW(p; r)$ is transient for every $r \in (0,1)$. The nature of the leading singularity of the Green's function $G(z)$, and its location *vis a vis* the singularity $z = \rho(p)$ of the first-passage generating function $F(z)$, depend on the value of r. Observe that if $p > 1/2$ then $F(1) < 1$, since the random walk is transient. In fact, by (6.6),

$$(6.13) \qquad F(\rho(p)) = F(1/(2\sqrt{p(1-p)})) = \sqrt{(1-p)/p} < 1,$$

and so, by the strict monotonicity of $F(z)$ on the positive real axis, $F(z) < 1$ for all $z \leq \rho(p)$.

Define

$$(6.14) \qquad r_* = r_*(p) = 1 - \frac{\sqrt{1-p}}{\sqrt{p} - \sqrt{1-p}}$$

LEMMA 6.1. *Let $z = \zeta = \zeta(p;r)$ be the smallest root of the equation*

$$(6.15) \qquad 1 = z(1-r) + rzF(z)$$

in the interval $z \in (0, \rho(p)]$, provided there is one. Then for each $p > 1/2$, the zero $\zeta(p;r)$ is continuous and strictly increasing in r for $0 < r \leq r_*(p)$. At $r = r_*(p)$,

$$(6.16) \qquad \zeta(p;r) = \rho(p) = 1/(2\sqrt{p(1-p)}).$$

For $r > r_*(p)$, there are no zeros of the equation (6.15) in the interval $(0, \rho(p)]$.

PROOF. Since $F(1) < 1$, the zero $\zeta(p; r)$, if it exists, must be strictly greater than 1. Moreover, since $F(z)$ is strictly increasing in $(0, \rho]$, there is at most one zero in $(0, \rho(p)]$. For each $z \in (0, \rho(p)]$, the function $z(1 - r) + rzF(z)$ is strictly decreasing in r, since $F(z) < F(\rho(p)) < 1$ by (6.13). Consequently, the function $\zeta(p; r)$ is strictly increasing in r. That $\zeta(p; r)$ is continuous in r follows from the implicit function theorem, as $F'(z) > 0$ for all $z \in (0, \rho(p)]$.

For all r such that $(1 - r)^{-1} \leq \rho(p)$, the function $z \mapsto 1 - z(1 - r) - rzF(z)$ must have a zero in the interval $(1, (1 - r)^{-1})$, by the intermediate value theorem. Hence, $\zeta(p; r)$ is defined for all small $r > 0$. By the preceding paragraph, the function $\zeta(p; r)$ is strictly increasing and continuous in r up to the point r_* where $\zeta(p; r_*) = \rho = \rho(p)$. This occurs at the solution $r = r_*$ of the linear equation

$$1 = \rho(1 - r) + \rho r F(\rho).$$

Using (6.6), it is easily checked that $r_* = r_*(p)$ is given by (6.14). $\square$

COLLARY 6.2. *Assume that $p > 1/2$. Let $r_* = r_*(p)$, $\zeta = \zeta(p; r)$, and $\rho = \rho(p)$ be as in Lemma 6.1 and equation (6.7). Then for the reflecting random walk $RRW(p; r)$ the transition probabilities satisfy*

$$(6.17) \qquad P^0\{X_n = 0\} \sim C\zeta^{-n} \qquad\qquad \text{if} \qquad 0 < r < r_*;$$

$$(6.18) \qquad P^0\{X_n = 0\} \sim C\rho^{-n} n^{-1/2} \qquad \text{if} \qquad r = r_*;$$

$$(6.19) \qquad P^0\{X_n = 0\} \sim (C + (-1)^n C')\rho^{-n} n^{-3/2} \qquad \text{if} \qquad r_* < r < 1,$$

for suitable constants $C = C(p; r) > 0$ and $C' = C'(p; r) \neq 0$.

PROOF. If $r < r_*(p)$ then by Lemma 6.1 the smallest positive zero $\zeta(p; r)$ of the equation (6.15) is strictly less than the radius $\rho(p)$ of convergence of $F(z)$. Hence, by equation (6.4), the smallest positive singularity of $G(z)$ is at $\zeta(p; r) < \rho(p)$. Since $F'(z) > 0$ at $z = \zeta(p; r)$, and since $F(z)$ is analytic in a neighborhood of $\zeta(p; r)$ it follows that $G(z)$ has a simple pole at $z = \zeta(p; r)$. It is not difficult to show, using the explicit expressions (6.4) and (6.5), that $G(z)$ has no other singularities on the circle $|z| = \zeta(p; r)$. The relation (6.17) follows.

Consider next the case $r > r_*(p)$. Since $\rho(1 - r) + \rho r F(\rho)$ is strictly decreasing in r (recall that $F(\rho) < 1$, by (6.13)), and takes the value 1 at $r_*(p)$, it follows that, for $r > r_*(p)$,

$$1 > \rho(1 - r) + \rho r F(\rho).$$

Consequently, by equation (6.4), $G(z)$ is an analytic function of z and $F(z)$ in a neighborhood of any point $(\xi, F(\xi))$ where $|\xi| = \rho$, and has power series expansion

$$G(z) = G(\xi) + G(\xi)^2 \left((z - \xi)(1 - r + rF(\xi)) + \xi r(F(z) - F(\xi)) + \cdots\right) + \cdots$$

Since $F(z)$ has branch points of order two at $z = \pm\rho$, so does $G(z)$; and since $F(z)$ has no other singularities on the circle $|z| = \rho$, neither does $G(z)$. Thus, by Darboux' theorem, the asymptotic formula (6.19) holds.

Finally, consider the case $r = r_*(p)$. In this case the Green's function once again has singularities at $z = \pm\rho$, and at no other points on $|z| = \rho$; however, $G(z)$ becomes infinite as $z \to \rho$, but remains finite as $z \to -\rho$. The nature of the singularity at $z = +\rho$ may be deduced from the identity

$$1/G(z) = 1 - z(1 - r) - zrF(z);$$

this implies that, in a neighborhood of $z = \rho$,

$$G(z) = A(z)(\rho - z)^{-1/2} + B(z)$$

for suitable analytic function $A(z), B(z)$. Similarly, near $z = -\rho$ the function $G(z)$ satisfies

$$G(z) = C(z)(\rho - z)^{1/2} + D(z)$$

for suitable analytic function $C(z), D(z)$. Thus, by Darboux' theorem, (6.18) holds.
□

Remark 1. It may strike the reader as somewhat strange that the period-two oscillation in the transition probabilities occurs only for $r > r_*(p)$. Here is an explanation: Conditional on the event $X_n = 0$, the distribution of the total number N_n of returns to 0 before time n satisfies

$$N_n \xrightarrow{P} \infty \qquad \text{if } r \leq r_*(p), \text{ but}$$
$$N_n = O_P(1) \qquad \text{if } r > r_*(p)$$

as $n \to \infty$. Since the state 0 has a positive holding probability, the many returns to 0 destroy any aperiodicity when $r \leq r_*(p)$.

Remark 2. The random walks considered in this section do not obey the hypotheses of Theorems 2.5-2.6, and so by themselves do not prove that all the different cases of these theorems actually occur. However, it is a routine (if tedious) matter to obtain explicit formulas for the Green's functions of modified reflecting random walks with holding probability $\alpha > 0$ for all states $x \geq 1$, and to deduce from these expressions local limit theorems similar to those obtained above. The only difference in the results is that the period-two oscillation in formula (6.19) does not occur.

7. Analysis of the leading singularity

We return now to the study of random walks on regular languages in general. Recall from section 5 that the matrix $\mathbf{G}(z)$ of Green's functions $G_{xy}(z)$ is a rational function (5.6) of the generating functions $H_{ab,c}(z)$, and that these generating functions are interrelated by the system (5.10) of quadratic equations (5.9). As earlier, we denote by $\mathbf{H}$ the vector of nondegenerate first-passage generating functions $H_{ab,c}(z)$.

7.1. Radius of convergence of the system (5.10). In this paragraph we will show that all of the *nondegenerate* first-passage generating functions $H_{ab,c}(z)$ (that is, all those that occur as entries of the vector-valued function $\mathbf{H}(z)$) have the same radius of convergence, and that all have the same *singularity type* at $z = \rho$. In general, two algebraic functions $\varphi(z)$ and $\psi(z)$, both with singularities at $z = \zeta$, will be said to have the same *singularity type* at the point $z = \zeta$ if (a) either $\varphi(z)$ and $\psi(z)$ are both finite at $z = \zeta$, or are both infinite; and (b) the leading terms in the Puiseux expansions of $\varphi(z)$ and $\psi(z)$ at $z = \zeta$ have the same exponent $\alpha > 0$, that is,

$$\varphi(z) \sim C(\zeta - z)^{-\alpha} \qquad \text{and} \qquad \psi(z) \sim C'(\zeta - z)^{-\alpha} \qquad \qquad \text{or}$$
$$\varphi(\zeta) - \varphi(z) \sim C(\zeta - z)^{\alpha} \qquad \text{and} \qquad \psi(\zeta) - \psi(z) \sim C'(\zeta - z)^{\alpha}$$

for suitable constants $C, C' \neq 0$.

PROPOSITION 7.1. *All of the generating functions $H_{ab,c}(z)$ occurring as entries of $\mathbf{H}(z)$ have the same radius ρ of convergence, and all have the same singularity type at $z = \rho$.*

The proof will rely on an irreducibility property of the system (5.10) of quadratic equations governing the generating functions $H_{ab,c}(z)$. The system (5.10) asserts that the function $\mathbf{H}(z)$ is a fixed point of the mapping $\mathbf{F} \mapsto z\mathbf{Q}(\mathbf{F})$; consequently, $\mathbf{H}(z)$ is a fixed point of any iterate of this mapping. In particular,

$$(7.1) \qquad \mathbf{H}(z) = (z\mathbf{Q}) \circ (z\mathbf{Q}) \circ \cdots \circ (z\mathbf{Q})(\mathbf{H}(z)) := z^m \mathbf{Q}^{\circ m}(\mathbf{H}(z)),$$

where $\circ$ denotes functional composition. The following lemma concerning this hierarchy of polynomial equations is the key point in the analysis where the Irreducibility Assumptions 2.3 and 2.4 are used.

LEMMA 7.2. *There exists an integer $m \geq 1$ such that, in each component of $(z\mathbf{Q})^{\circ m}(\mathbf{H})$, each variable $H_{ab,c}$ appears in a term with* positive *coefficient.*

PROOF. In each component equation (5.9) of the vector equation (5.10), the right side contains a term that is z times a constant $p(ab|ab)$ times the left side. Consequently, when equation (5.10) is substituted in the mth iterate to obtain the $(m + 1)-$th iterate, any product of the variables $H_{ab,c}$ that occurred as a term in the mth iterate will remain as a term in the $(m+1)-$th iterate, albeit with a higher power of z. Therefore, to prove the lemma it suffices to show that, for each pair $H_{ab,c}, H_{a'b',c'}$ of components of $\mathbf{H}$, there exists $m \geq 1$ such that $H_{a'b',c'}$ appears as a factor in some term of the $H_{ab,c}-$component of the $m-$th iterate of (5.10).

Recall that if $H_{ab,c}(z), H_{a'b',c'}(z)$ are entries of $\mathbf{H}(z)$ then they are not identically zero, equivalently, there exist positive-probability paths from ab to c and from $a'b'$ to c' through words of length ≥ 2. Let

$$ab \to w_1 \to w_2 \to \cdots \to w_n \to c$$

be such a path from ab to c. By Assumption 2.4, there is a positive-probability path from w_n to a word $w_n wa'b'$ through the set of words prefaced by w_n. By the homogeneity property (5.8),

$$E^{w_n wa'b'} z^T \mathbf{1}\{X_T = w_n wc'\} = H_{a'b',c'}(z) \not\equiv 0.$$

By Assumption 2.3, for each consecutive pair β, γ of letters in the word wc', there is a positive-probability path from $\beta\gamma$ to β through words prefaced by β. Consequently, for sufficiently large m, the equation for $H_{ab,c}(z)$ in the mth iterate of the system (5.10) contains a term with the factor $H_{a'b',c'}(z)$. $\qquad\square$

Proof of Proposition 7.1. That all of the generating functions $H(z)$ have the same radius of convergence follows immediately from Lemma 7.2, because, since $H(1) > 0$ and $H'(z) > 0$ for $z > 0$ and for each function H, Lemma 7.2 implies that for any pair H, H_* there are constants $C, C' > 0$ such that

$$(7.4) \qquad\qquad H(z) \geq CH_*(z) \qquad \text{and} \qquad H_*(z) \geq C'H(z)$$

for all $z \geq 1$. It also follows from this that either all of the values $H(\rho)$ are $+\infty$, or all are finite (and, of course, positive). In the case where all are $+\infty$, the inequalities (7.4) imply that the leading terms in the Puiseux expansions all have the same exponent α.

Suppose that all of the functions H are finite (and positive) at their common radius of convergence $\rho \geq 1$. By Lemma 7.2, for each integer $k \geq 1$ there exist constants $C_k, C'_k > 0$ such that, for any pair H, H_*, the kth derivatives satisfy

$$(7.5) \qquad H^{(k)}(z) \geq C_k H_*^{(k)}(z) \qquad \text{and} \qquad H_*^{(k)}(z) \geq C'_k H^{(k)}(z).$$

It follows that the leading terms in the Puiseux expansions must all have the same exponent α.

Henceforth, we shall refer to ρ as the "radius of convergence" of the system (5.10).

LEMMA 7.3. *Let $H_{ab,c}(z)$ be nondegenerate, and let $0 < r \leq \rho$ be such that $H_{ab,c}(r) < \infty$. Then for any complex number $\zeta \neq r$ such that $|\zeta| = r$,*

$$(7.6) \qquad |H_{ab,c}(\zeta)| < H_{ab,c}(r)$$

PROOF. This is a consequence of Assumption 2.1, which implies that, for some $\varepsilon > 0$, $p(ab|ab) \geq \varepsilon$. This, in turn, implies that, for each integer $n \geq 1$,

$$(7.7) \qquad P^{ab}\{T = n+1; X_{n+1} = c\} \geq \varepsilon P^{ab}\{T = n; X_n = c\}.$$

Consider any two two consecutive terms of the power series expansion of $H_{ab,c}(\zeta)$, with $a_n := P^{ab}\{T = n; X_n = c\}$:

$$\begin{aligned}
|a_n \zeta^n + a_{n+1}\zeta^{n+1}| &= r^n |a_n + a_{n+1}\zeta| \\
&< r^n |a_n + a_{n+1}r| \\
&= a_n r^n + a_{n+1}r^{n+1}.
\end{aligned}$$

This shows that the sum of any two consecutive nonzero terms in the power series expansion of $H_{ab,c}(\zeta)$ is strictly less, in modulus, than the sum of the corresponding terms in the power series expansion of $H_{ab,c}(r)$. □

7.2. Radius of convergence of the Green's functions $G_{xy}(z)$.

PROPOSITION 7.4. *All of the Green's functions $G_{xy}(z)$ have the same radius of convergence $R \geq 1$. Moreover, all of the Green's functions $G_{xy}(z)$ have the same singularity type at $z = R$.*

PROOF. The argument follows similar lines as the proof of Proposition 7.1, but here Lemma 7.2 is replaced by a "Harnack"-type inequality. By Assumption 2.2, for any two elements u, v of the regular language $\mathcal{L}$, there are positive probability paths from u to v and from v to u. Consequently, for any two pairs $(x, y), (x', y') \in \mathcal{L}^2$ there exist $k \geq 0$ and $\varepsilon > 0$ such that

$$(7.8) \qquad P^x\{X_{n+k} = y\} \geq \varepsilon P^{x'}\{X_n = y'\}.$$

This inequality clearly implies that the Green's functions $G_{xy}(z)$ and $G_{x'y'}(z)$ have the same radius of convergence R, and that if $G_{x'y'}(R) = \infty$ then $G_{xy}(R) = \infty$. It follows that the Green's functions $G_{xy}(z)$ are either all finite or all infinite at their common radius of convergence R.

Because each $G_{xy}(z)$ is an algebraic function, it has a Puiseux expansion at $z = R$. Since $G_{xy}(z)$ is strictly increasing in z along the line segment $[0, R]$, the

first two terms of the Puiseux series must be of the form

$$(7.9) \qquad G_{xy}(R) - G_{xy}(z) \sim C_{xy}(R-z)^{\alpha_{xy}} \qquad \text{if } G_{xy}(R) < \infty; \qquad \text{or}$$

$$(7.10) \qquad\qquad G_{xy}(z) \sim C_{xy}(R-z)^{-\alpha_{xy}} \qquad \text{if } G_{xy}(R) = \infty,$$

where $\alpha_{xy} > 0$ and $C_{xy} > 0$. Consider first the case where all the Green's functions are infinite at $z = R$: By inequality (7.8) above, for any two pairs $(x,y),(x',y') \in \mathcal{L}$, and any $\alpha > 0$, there exists a constant $\varepsilon > 0$ such that

$$(R-z)^\alpha G_{xy}(z) \geq \varepsilon (R-z)^\alpha G_{x'y'}(z).$$

Consequently, in case (7.10), all of the exponents α_{xy} must be the same.

Now consider the case (7.9). Once again, by inequality (7.8), for any two pairs $(x,y),(x',y') \in \mathcal{L}$, and any $\alpha > 0$, there exist constants $\varepsilon_k > 0$ such that the kth derivatives of the Green's functions satisfy

$$(R-z)^\alpha G_{xy}^{(k)}(z) \geq \varepsilon_k (R-z)^\alpha G_{x'y'}^{(k)}(z).$$

This implies that in case (7.9) all of the exponents α_{xy} are the same. $\qquad\square$

7.3. The Perron-Frobenius matrices $M(z) = (\mathbf{P} + \mathbf{K}(z))$. It is not necessarily the case that $R = \rho$: as we have seen in section 6, it is possible that $R < \rho$, and also that $R = \rho$. Later we will show that $R > \rho$ is impossible. Recall equation (5.6), according to which

$$(7.13) \qquad\qquad \mathbf{I} = (\mathbf{I} - z\mathbf{P} - z\mathbf{K}(z))\, \mathbf{G}(z).$$

This implies that if z is such that the matrix $(\mathbf{I} - z\mathbf{P} - z\mathbf{K}(z))$ is invertible, and if $\mathbf{K}$ is analytic in a neighborhood of z, then the entries G_{xy} of $\mathbf{G}$ are finite and analytic near z, because (5.7) implies that they may be expressed as convergent Neumann series in the functions K_{xy}. Consequently, if $z = \zeta$ is a singularity of one of the Green's functions $G_{xy}(z)$ then either $z = \zeta$ is a singularity of $\mathbf{H}(z)$ or the matrix $(\mathbf{I} - z\mathbf{P} - z\mathbf{K}(z))$ is *noninvertible*. The following result will be of crucial importance in investigating the invertibility/noninvertibility of this matrix.

LEMMA 7.5. *For $0 < z \leq \rho$ the matrix $M(z) := \mathbf{P} + z\mathbf{K}(z)$ is a Perron-Frobenius matrix.*

PROOF. Recall that a Perron-Frobenius matrix is a square matrix M such that some positive power M^n of M has strictly positive entries. For $z > 0$, the matrices $\mathbf{P}$ and $\mathbf{K}(z)$ have entries

$$(7.14) \qquad\qquad \mathbf{P}_{xy} := p(y|x),$$

$$(7.15) \qquad\qquad K_{xy}(z) := \sum_{a,b} p(ab|x) H_{ab,y}(z),$$

where $p(\cdot|\cdot)$ are the transition probabilities in (2.1) and $H_{ab,c}(z)$ is defined by (5.2). The indices x,y are words in $\mathcal{L}$ of length 0 or 1. For any such pair x,y of indices, there is a positive-probability path from x to y, by Assumption 2.2. If there is a positive-probability path of length $n \geq 1$ through words of length ≤ 1, then certainly $\mathbf{P}_{xy}^n > 0$; and since $p(w|w) > 0$ for all elements of $\mathcal{L}$, by Assumption 2.1, it then follows that $\mathbf{P}_{xy}^{n+k} > 0$ for all $k \geq 0$. If there is *not* a positive-probability path from x to y through words of length ≤ 1, then there must be a positive-probability probability path from x to y that visits a word of length 2. Such a path visits a sequence of words of length ≤ 1, then a word w of length 2, then, after wandering

among the words of length ≥ 2, another sequence of words of length 1, etc., and finally arrives at y:

$$x \rightarrow x_1 \rightarrow x_2 \rightarrow \cdots x_{k_1} \rightarrow w_1$$
$$\rightarrow x_{k_1+1} \rightarrow x_{k_1+2} \rightarrow \cdots \rightarrow x_{k_2} \rightarrow w_2$$
$$\rightarrow x_{k_2+1} \rightarrow \cdots \rightarrow \cdots$$
$$\rightarrow x_{k_r} \rightarrow w_r$$
$$\rightarrow x_m = y$$

For each pair $w = w_i, v = x_{k_i+1}$ in this sequence, it must be the case that $H_{w,v}(z) > 0$ for $z > 0$, because there is a positive-probability path from w to v through words of length ≥ 2. Moreover, for each pair x_i, x_{i+1} not straddling a w_i, it must be the case that $p(x_{i+1}|x_i) > 0$. Consequently, for all $z > 0$,

$$(\mathbf{P} + \mathbf{K}(z))_{xy}^{r+m} > 0 \qquad \Longrightarrow$$
$$(\mathbf{P} + \mathbf{K}(z))_{xy}^{r+m+k} > 0 \qquad \forall\, k \geq 0.$$

It now follows that, for all sufficiently large $n \geq 1$, the matrix $(\mathbf{P} + \mathbf{K}(z))^n$ has strictly positive entries. $\qquad\square$

Because $M(z) := (\mathbf{P} + \mathbf{K}(z))$ is a Perron-Frobenius matrix for all $0 < z < \rho$, the conclusions of the Perron-Frobenius theorem (see, for example, [**21**]) apply: In particular, $M(z)$ has a positive eigenvalue $\mu = \mu_z = \mu(z)$ of multiplicity one, and the rest of its spectrum is contained in the open disc $|\zeta| < \mu_z$. Moreover, there are unique right and left eigenvectors $\alpha = \alpha_z = \alpha(z)$ and $\beta = \beta_z = \beta(z)$, with *strictly* positive entries, normalized so that the following system of equations hold:

(7.16)
$$M(z)\beta(z) = \mu(z)\beta(z); \qquad \beta(z)^T \alpha(z) = 1;$$
$$\alpha(z)^T M(z) = \mu(z)\alpha(z)^T; \qquad \alpha(z)^T \mathbf{1} = 1,$$

where, in the last equation, $\mathbf{1}$ denotes the vector with all entries 1 (thus, the equation asserts that $\alpha(z)$ is a probability vector for each $z \in (0, \rho)$).

LEMMA 7.6. *The functions $\mu(z)$, $\alpha(z)$, and $\beta(z)$ extend to analytic functions in a neighborhood of the line segment $(0, \rho)$, and for $z \in (0, \rho)$, the function $\mu(z)$ is nondecreasing in z. Furthermore, if all of the first-passage generating functions $H_{ab,c}(z)$ are finite at $z = \rho$, then $\mu(z)$, $\alpha(z)$, and $\beta(z)$ are continuous on the* closed *interval $z \in [0, \rho]$.*

PROOF. The entries of $M(z)$, being linear combinations of the first-passage generating functions $H_{ab,c}(z)$, are analytic functions of z for $|z| < \rho$. Moreover, if all of the functions $H_{ab,c}(z)$ are finite at $z = \rho$, then as $z \rightarrow \rho-$ the entries of $M(z)$ converge monotonically to those of $M(R)$. Consequently, the analyticity and continuity of $\mu(z)$, $\alpha(z)$, and $\beta(z)$ follow from standard results in regular perturbation theory (see, for example, [**19**], section XII.1), because the lead eigenvalue $\mu(z)$ has multiplicity one at every $z \in (0, \rho)$. That the eigenvalue $\mu(z)$ is nondecreasing in z for $0 \leq z \leq \rho$ follows because the entries of $M(z)$ are nondecreasing in z. $\qquad\square$

LEMMA 7.7. $z\mu(z) < 1$ *for all $z \in [0, R \wedge \rho)$.*

PROOF. Suppose not; then there exists $\zeta \in (0, R \wedge \rho)$ such that $\mu(\zeta) = 1$. By Lemma 7.6, the function $z\mu(z)$ is strictly increasing in z along the line segment $[0, \rho)$, and so $z\mu(z) < 1$ for all $z \in [0, \zeta)$. Because $z\mu(z)$ is the spectral radius of $zM(z)$, it follows that the matrix $(\mathbf{I} - zM(z))$ is invertible for all $z \in (0, \zeta)$. Consequently, the equation (7.13) is solvable for $\mathbf{G}(z)$ along $(0, \zeta)$, via the equation (5.7), and so, by the Law of Permanence of functional equations, the relation (7.13) holds for all z in a neighborhood of the segment $[0, \zeta)$. Taking the inner products with the left eigenvector $\alpha(z)^T$ and the vector $\mathbf{1}$ now gives

$$(7.17) \qquad\qquad 1 = (1 - z\mu(z))\alpha(z)^T \mathbf{G}(z)\mathbf{1},$$

valid for all $z \in [0, \zeta)$. Now $\zeta < R \wedge \rho$, by hypothesis, and so each of the functions $\mu(z), \alpha(z)$, and $\mathbf{G}(z)$ is continuous at $z = \zeta$. Thus, relation (7.17) persists at $z = \zeta$. But $\zeta\mu(\zeta) = 1$, and so equation (7.17) cannot possibly hold at $z = \zeta$ unless $G(\zeta) = \infty$, which would imply that $\zeta = R$. $\square$

LEMMA 7.8. *For each $r \in (0, \rho)$ and each $\zeta \neq r$ of modulus $|\zeta| = r$, the spectral radius of $M(\zeta)$ is strictly less than $\mu(r)$.*

PROOF. The entries of $M(z) = (\mathbf{P} + \mathbf{K}(z))$ are given by equations (7.14) and (7.15). The nonzero entries of $\mathbf{K}(z)$ are positive linear combinations of the functions $H_{ab,c}(z)$. By Lemma 7.3, these satisfy the inequalities

$$|H_{ab,c}(\zeta)| < H_{ab,c}(r).$$

Consequently, the nonzero entries of $M(\zeta)$ are strictly less, in modulus, than the corresponding entries of $M(r)$. Therefore, the spectral radius of $M(\zeta)$ is strictly less than that of $M(r)$. $\square$

PROPOSITION 7.9. *If $R < \rho$ then $z = R$ is a simple pole of each $G_{xy}(z)$.*

PROOF. Recall that R, being the common radius of convergence of the Green's functions $G_{xy}(z)$, is a singularity of $\mathbf{G}(z)$. Recall also that the functions $H_{ab,c}(z)$ are analytic in the disc $|z| < \rho$, which contains the closed disc $|z| \leq R$. Hence, if $R < \rho$, then, by equation (7.13), the matrix $(\mathbf{I} - M(z))$ must be noninvertible at $z = R$, because if it were invertible then the function $\mathbf{G}(z)$ would extend analytically to a neighborhood of $z = R$, which it does not.

By the preceding lemma, the Perron-Frobenius eigenvalue $\mu(z)$ of $M(z)$ is strictly less than $1/z$ for all $z \in (0, R)$, and, by Lemma 7.6, it is continuous in z for $0 \leq z < \rho$. Since $\mu(z)$ is the spectral radius of $M(z)$, if $R\mu(R) < 1$ then the matrix $(\mathbf{I} - RM(R))$ would be invertible, which it is not. Therefore, if $R < \rho$ then

$$(7.19) \qquad\qquad\qquad R\mu(R) = 1.$$

Now consider the function $F(z) := \alpha(z)^T \mathbf{G}(z)\mathbf{1}$. This is a linear combination of the entries $G_{xy}(z)$ of the Green's matrix $\mathbf{G}(z)$. These entries are multiplied by entries of the left eigenvector $\alpha(z)^T$, which are strictly positive for $0 < z < \rho$ and

analytic in z. By equation (5.6), for all $z \in (0, R)$,

$$(7.20) \qquad \begin{aligned} F(z) &= \alpha(z)^T \mathbf{G}(z)\mathbf{1} \\ &= \alpha(z)^T \mathbf{1} + z\alpha(z)^T M(z)\mathbf{G}(z)\mathbf{1} \\ &= 1 + z\mu(z)F(z) \\ &= (1 - z\mu(z))^{-1}, \end{aligned}$$

which shows that $F(z)$ has a simple pole at $z = R$, since $z\mu(z) \to 1$ as $z \to R$ and $\mu(z)$ is nondecreasing in z. Since $F(z)$ is a positive linear combination of the entries of the Green's function matrix $\mathbf{G}(z)$, and since all of these entries have the same singularity type at $z = R$ (by Proposition 7.4) it follows that that each of these entries has a simple pole at $z = R$. $\qquad \square$

COLLARY 7.10. *If $R < \rho$ then for each pair $x, y \in \mathcal{L}$ of words of length 0 or 1, there exists a constant $C_{xy} > 0$ such that, as $n \to \infty$,*

$$(7.21) \qquad P^x\{X_n = y\} \sim C_{xy}R^{-n};$$

PROOF. By Proposition 7.9, if $R < \rho$ then $G(z)$ has a simple pole at $z = R$. By Lemma 7.8, if $\zeta \neq R$ and $|\zeta| = R$ then the spectral radius of $M(\zeta)$is strictly less than that of $M(R)$, which, by equation (7.19), is $1/R$. Consequently, for each such ζ, the matrix $(\mathbf{I} - \zeta M(\zeta))$ is invertible. Since the entries of $M(z)$ are analytic in $|z| < \rho$, it follows by the implicit function theorem that, for each $\zeta \neq R$ such that $|\zeta| = R$, the Green's function matrix $\mathbf{G}(z)$ is analytic in a neighborhood of $z = \zeta$. Thus, $\mathbf{G}(z)$ has no other singularities in the closed disc $|z| \leq R$. It follows that the coefficients $P^x\{X_n = y\}$ of $G_{xy}(z)$ must obey the asymptotic law (7.21) for some nonzero constant C_{xy}. Since the coefficients are probabilities, the constant must be positive. $\qquad \square$

PROPOSITION 7.11. *If $R \not< \rho$, then $R = \rho$. In this case, for each entry $H_{ab,c}$ of* **H**,

$$(7.22) \qquad H_{ab,c}(\rho) < \infty.$$

PROOF. By Lemma 7.7, if $R \geq \rho$ then $z\mu(z) < 1$ for all $z \in [0, \rho)$. It follows that the entries of the matrix $zM(z)$ remain bounded as $z \to \rho-$. Since the entries of $M(z)$ are positive linear combinations of the probabilities $p(y|x)$ and the functions $H_{\beta,c}(z)$ it follows that at least some of the functions $H_{\beta,c}(z)$ remain bounded as $z \to \rho-$. Proposition 7.1 implies that *all* of the functions $H_{\beta,c}(z)$ remain bounded as $z \to \rho-$. $\qquad \square$

7.4. The Perron-Frobenius matrix of the system (5.10). Having established the local limit theorem (7.21) in the case $R < \rho$, we now turn our attention to the case

$$(7.23) \qquad R = \rho.$$

By Proposition 7.9, if (7.23) holds then all of the entries $H_{ab,c}(z)$ of $\mathbf{H}(z)$ are finite at $z = \rho$. Since all of the generating functions $H_{ab,c}(z)$ occurring as entries of $\mathbf{H}$ have singularities at $z = \rho$, where ρ is the common radius of convergence, the point $z = \rho$ must be a branch point of the algebraic curve (5.10). In particular, by the implicit function theorem, $z = \rho$ must be the smallest positive real number

where the *linearization* of the system (5.10) becomes singular. In vector form, the system (5.10) reads $\mathbf{H}(z) = z\mathbf{Q}(\mathbf{H}(z))$; the linearization is obtained by taking the differential of each side:

$$d\mathbf{H} = (dz)\mathbf{Q}(\mathbf{H}) + \left(\frac{\partial\mathbf{Q}}{\partial\mathbf{H}}\right)(d\mathbf{H}) \qquad \Longleftrightarrow$$

(7.25) $$\left(\mathbf{I} - z\frac{\partial\mathbf{Q}}{\partial\mathbf{H}}\right)(d\mathbf{H}) = (dz)\mathbf{Q}(\mathbf{H})$$

The analysis of the linear system (7.25) is similar to the analysis of the linear system (7.13) in paragraph 7.3 above. Once again, a crucial role is played by a one-parameter family of matrices which are Perron-Frobenius matrices along the line segment $(0, \rho]$. These are defined as follows:

(7.26) $$\mathbf{L}_z = \mathbf{L}(z) = \left(\frac{\partial\mathbf{Q}}{\partial\mathbf{H}}\right)_{\mathbf{H}=\mathbf{H}(z)}$$

LEMMA 7.12. *For each $z \in (0, \rho]$, the matrix $\mathbf{L}_z$ is a Perron-Frobenius matrix.*

PROOF. Since the components of $\mathbf{Q}(\mathbf{H})$ are polynomials in the variables $H_{ab,c}$ with positive coefficients, so are the entries of the matrix $\mathbf{L}_z = \partial\mathbf{Q}/\partial\mathbf{H}$, as are the entries of any positive power $\mathbf{L}_z^m$ of $\mathbf{L}_z$. Since those generating functions $H_{ab,c}(z)$ that occur as entries of the vector $\mathbf{H}(z)$ are strictly positive for $0 < z \leq \rho$, it follows that for each $z \in (0, \rho]$ and $m \geq 1$, the entries of $\mathbf{L}_z^m$ are nonnegative. It remains to be shown that, for sufficiently large m, these entries are nonzero.

Let $H = H_{ab,c}$ and $H_* = H_{a'b',c'}$ be entries of $\mathbf{H}$, and, for each $m \geq 1$, let $Q_H^{\circ m}(\mathbf{H})$ be the $H-$component of the vector polynomial $\mathbf{Q}^{\circ m}(\mathbf{H})$. We will show that the $(H, H_*)-$entry of the matrix $\mathbf{L}_z^m$ is nonzero if and only if the variable H_* occurs as a factor in a term of $Q_H^{\circ m}(\mathbf{H})$. The proof is by induction on m. When $m = 1$, the statement is certainly true, because $\mathbf{L}_z = \partial\mathbf{Q}/\partial\mathbf{H}$. Assume, then that the claim is true for all positive integers $\leq m$. Consider the $(H, H_*)-$entry of the matrix $\mathbf{L}_z^{m+1}$; this will be nonzero if and only if, for some entry H_{**} of $\mathbf{H}$, the $(H, H_{**})-$entry of $\mathbf{L}_z^m$ and the $(H_{**}, H_*)-$entry of $\mathbf{L}_z$ are both nonzero. By the induction hypothesis, this occurs if and only if H_{**} occurs as a factor in a term of $Q_H^{\circ m}(\mathbf{H})$ and H_* occurs as a factor in a term of $Q_H(\mathbf{H})$. But this occurs if and only if H_* occurs as a factor in a term of $Q_H^{\circ(m+1)}(\mathbf{H})$, because the vector polynomial $\mathbf{Q}^{\circ(m+1)}(\mathbf{H})$ is obtained by substituting $\mathbf{Q}(\mathbf{H})$ for $\mathbf{H}$ in $\mathbf{Q}^{\circ m}(\mathbf{H})$.

Now recall Lemma 7.2, according to which there exists $m \geq 1$ such that, in each component of $\mathbf{Q}^{\circ m}$, *every* entry H of $\mathbf{H}$ occurs as a factor in a term with positive coefficient. In view of the result of the previous paragraph, it follows that, for any such m, every entry of $\mathbf{L}_z^m$ is nonzero.   □

Remark. The argument used in the second paragraph of the proof may be modified to prove a somewhat stronger statement. If $m \geq 1$ is so large that every component of $\mathbf{Q}^{\circ m}(\mathbf{H})$ contains every $H = H_{ab,c}$ as a factor in a nonzero term, then, by resubstitution, it follows that every component of $\mathbf{Q}^{\circ 2m}(\mathbf{H})$ contains the *square* of every $H = H_{ab,c}$ as a factor in a nonzero term. Consequently, every entry of the matrix $\mathbf{L}_z^{2m}$ contains every $H = H_{ab,c}$ as a factor in a nonzero term.   □

Because $\mathbf{L}_z$ is a Perron-Frobenius matrix when $0 < z \leq \rho$, it is governed by the conclusions of the Perron-Frobenius theorem. Thus, it has a positive eigenvalue

$\lambda = \lambda_z = \lambda(z)$ of multiplicity one, and the rest of its spectrum is contained in the open disc $|\zeta| < \lambda_z$. Moreover, there are unique right and left eigenvectors $\mathbf{w} = \mathbf{w}_z = \mathbf{w}(z)$ and $\mathbf{v} = \mathbf{v}_z = \mathbf{v}(z)$, with *strictly* positive entries, satisfying

$$(7.27) \qquad \mathbf{L}(z)\mathbf{v}(z) = \lambda(z)\mathbf{v}(z); \qquad \mathbf{v}(z)^T\mathbf{w}(z) = 1;$$
$$\mathbf{w}(z)^T\mathbf{L}(z) = \lambda(z)\mathbf{w}(z)^T; \qquad \mathbf{w}(z)^T\mathbf{1} = 1,$$

where, in the last equation, $\mathbf{1}$ denotes the vector with all entries 1. Since the matrix $\mathbf{L}(z)$ depends analytically on z for $|z| < \rho$, and since the multiplicity of the eigenvalue $\lambda(z)$ is 1 at every $z \in (0, \rho]$, the functions $\lambda(z), \mathbf{w}(z)$, and $\mathbf{v}(z)$ are all analytic functions of z for z in some open neighborhood of the line segment $(0, \rho]$ (see [**19**], section XII.1). Since the entries of the matrices $\mathbf{L}_z$ are nondecreasing in z, for $0 \le z \le \rho$, the Perron-Frobenius eigenvalue $\lambda(z)$ is also nondecreasing in z.

LEMMA 7.13. $\lambda(\rho) = 1/\rho$.

PROOF. Suppose that $\rho\lambda(\rho) < 1$. Since the spectrum of $\rho\mathbf{L}_\rho$ lies entirely in the disc $|\zeta| \le \rho\lambda(\rho)$, the matrix $I - \rho\mathbf{L}_\rho$ would be invertible. Consequently, the linearization (7.25) of the algebraic system (5.10)) at $z = \rho$ would be solvable for $d\mathbf{H}$:

$$d\mathbf{H} = dz\,(\mathbf{I} - z\mathbf{L}_z)^{-1}\,z\mathbf{Q}(\mathbf{H})).$$

But this would imply, by the implicit function theorem, that $\mathbf{H}(z)$ admits an analytic continuation to a neighborhood of $z = \rho$, which is impossible, since $z = \rho$ is a singularity of $\mathbf{H}(z)$.

Now suppose that $\rho\lambda(\rho) > 1$. Since $z \mapsto z\lambda(z)$ is analytic, and therefore continuous, along the line segment $(0, \rho]$ and takes the value 0 at $z = 0$, it would follow, by the intermediate value theorem, that $r\lambda(r) = 1$ for some $r \in (0, \rho]$. But this would imply that the matrix $(I - r\mathbf{L}_r)$ is singular, and that the system (7.25) cannot be solved for $d\mathbf{H}$ in terms of dz at $z = r$. This, in turn, would imply that $z = r$ is a singularity of the function $\mathbf{H}(z)$, a contradiction, since $z = \rho$ is the smallest positive singularity. $\qquad\square$

COLLARY 7.14. *If the RWRL X_n is transient, then $\rho > 1$.*

PROOF. If the RWRL is transient, then at least one of the first-passage generating functions $H_{ab,c}(z)$ must satisfy

$$(7.29) \qquad\qquad H_{ab,c}(1) < 1.$$

$\qquad\square$

LEMMA 7.15. *For each $z \ne \rho$ in the disc $|z| \le \rho$, the spectral radius of the matrix L_z is strictly less than $1/\rho$.*

PROOF. It suffices to show that, for some integer $m \ge 1$, the spectral radius of $\mathbf{L}_z^m$ is less than $1/\rho^m$. Recall that the entries of the matrix $\mathbf{L}_z^m$ are positive linear combinations of products of the variables z and $H_{ab,c}$. Moreover, if m is sufficiently large, then each entry contains every $H_{ab,c}$ as a factor in some term with positive coefficient (see the remark following the proof of Lemma 7.12). Consequently, by Lemma 7.3, if $z \ne \rho$ but $|z| = \rho$ then all of the entries of $\mathbf{L}_z^m$ are strictly less in modulus than the corresponding entries of $\mathbf{L}_\rho^m$. It now follows by a routine argument that the spectral radius of $\mathbf{L}_z^m$ is strictly less than that of $\mathbf{L}_\rho^m$, which, by Lemma 7.13, is $1/\rho^m$. $\qquad\square$

LEMMA 7.16. *None of the functions $H(z)$ occurring as entries of $\mathbf{H}$ have singularities on the circle $|z| = \rho$ other than those at $z = \rho$.*

PROOF. By Lemma 7.15, if $z \neq \rho$ but $|z| = \rho$ then the spectral radius of $z\mathbf{L}_z$ is strictly less than 1, and so the matrix $\mathbf{I} - z\mathbf{L}_z$ is invertible. Thus, the linearization (7.25) of the system (5.10) is solvable for $d\mathbf{H}$ in terms of dz at z. By the implicit function theorem, $\mathbf{H}$ admits an analytic continuation to a neighborhood of z, and so z is a regular point of $\mathbf{H}$. □

It remains to determine the nature of the singularity at $z = \rho$. By Corollary 5.2, each of the generating functions $H_{ab,c}(z)$ is algebraic; by Proposition 7.1, that all of the *nondegenerate* $H_{ab,c}(z)$ have the same singularity type at $z = \rho$; and by Proposition 7.11, if $\rho = R$ then all of the functions $H_{ab,c}(z)$ are finite at $z = \rho$. Therefore, there exist constants $\alpha > 0$ and $C_H > 0$ such that for each nondegenerate $H = H_{ab,c}$,

$$(7.30) \qquad\qquad H(\rho) - H(z) \sim C_H(\rho - z)^{\alpha}.$$

PROPOSITION 7.17. $\alpha = 1/2$.

PROOF. Consider once again the system $\mathbf{H}(z) = z\mathbf{Q}(\mathbf{H}(z))$ of quadratic equations (5.9) for the vector $\mathbf{H}(z)$ of nondegenerate first-passage generating functions $H_{ab,c}(z)$. These equations are valid for all z in the disc $|z| < \rho$, and, by continuity, for all z on the circle of convergence $|z| = \rho$. Taking differences, we have

$$(7.31) \qquad \mathbf{H}(\rho) - \mathbf{H}(z) = \rho\mathbf{Q}(\mathbf{H}(\rho)) - z\mathbf{Q}(\mathbf{H}(z))$$
$$= \rho(\mathbf{Q}(\mathbf{H}(\rho)) - \mathbf{Q}(\mathbf{H}(z))) + (\rho - z)\mathbf{Q}(\mathbf{H}(\rho))$$
$$- (\rho - z)(\mathbf{Q}(\mathbf{H}(\rho)) - \mathbf{Q}(\mathbf{H}(z))).$$

Writing $\Delta\mathbf{H} := \mathbf{H}(\rho) - \mathbf{H}(z)$ and $\Delta z := \rho - z$ and collecting terms, we obtain

$$(7.32) \qquad (\mathbf{I} - \rho\mathbf{L}_\rho)\,\Delta\mathbf{H} = (\Delta z)\mathbf{Q}(\mathbf{H}(\rho))$$
$$- \rho(\text{quadratic terms in } \Delta\mathbf{H})$$
$$- (\Delta z)(\text{linear and quadratic terms in } \Delta\mathbf{H}).$$

Since the quadratic polynomials $\mathbf{Q}(\mathbf{H})$ have nonnegative coefficients, so do the quadratic terms on the right side of (7.32), and at least some of these coefficients are positive. Now take the inner product with the left eigenvector $\mathbf{v}_\rho$ of the Perron-Frobenius matrix $\mathbf{L}_\rho$: since $\rho\lambda(\rho) = 1$, by Lemma 7.13, this gives

$$(7.33) \qquad (\Delta z)(\mathbf{v}_\rho^T\mathbf{Q}(\mathbf{H}(\rho))) = \rho(\text{quadratic form in } \Delta\mathbf{H})$$
$$+ (\Delta z)(\text{linear and quadratic terms in } \Delta\mathbf{H}).$$

Since the eigenvector $\mathbf{v}_R$ has strictly positive entries, the coefficient $\mathbf{v}_\rho^T\mathbf{Q}(\mathbf{H}(\rho))$ of Δz on the left side is positive. Moreover, the coefficients in the quadratic form are nonnegative, and at least some of them are positive; hence, the quadratic form is nonnegative definite and nondegenerate. This implies that the exponent in (7.30) must be $\alpha = 1/2$. □

COLLARY 7.18. *Assume that $R = \rho$. Then there exists $C_\mu > 0$ such that as $z \to R-$,*

$$(7.34) \qquad\qquad \mu(R) - \mu(z) \sim C_\mu(R - z)^{1/2}.$$

PROOF. When $R = \rho$, the first-passage generating functions $H_{ab,c}(z)$ are finite at $z = R$, by Proposition 7.11. Consequently, the entries of $M(z)$ are finite at $z = R$, because they are positive linear combinations of the functions $H_{ab,c}(z)$. Since the matrix $M(R)$ is a Perron-Frobenius matrix, its leading eigenvalue has multiplicity one, and so, by regular perturbation theory [**19**], the leading eigenvalue is an analytic function of the matrix entries in a neighborhood of $M(R)$. Thus, in particular, the function $\mu(z)$ is given by a convergent power series in the entries of $M(R) - M(z)$, and the linear terms of this power series are not all zero (because locally, near $M(R)$, the spectral radius must increase linearly with any entry). The relation (7.34) therefore follows from Proposition 7.17. $\square$

7.5. Behavior of the Green's functions at $R = \rho$. When $R = \rho$, all of the first-passage generating functions $H_{ab,c}(z)$ are finite at $z = R$, by Proposition 7.11, have Puiseux series of the form (7.30) at R, and have no other singularities on the circle $|z| = R$, by Lemma 7.16. By Lemma 7.7, for every $z \in (0, R)$ the spectral radius $z\mu(z)$ of $zM(z) := z(\mathbf{P} + \mathbf{K}(z))$ is strictly less than one, and so by the continuity of $\mu(z)$, either

$$(7.35) \qquad\qquad R\mu(R) < 1 \qquad \text{or}$$

$$(7.36) \qquad\qquad R\mu(R) = 1$$

The special cases considered in section 6 show that both possibilities may occur.

PROPOSITION 7.19. *For each entry $G_{xy}(z)$ of the Green's function matrix $\mathbf{G}(z)$, there exists a constant $C_{xy} > 0$ such that, as $z \to R-$,*

$$(7.37) \qquad\qquad G_{xy}(z) \sim C_{xy}(R - z)^{-1/2} \qquad \text{if } R\mu(R) = 1; \text{ and}$$

$$(7.38) \qquad G_{xy}(R) - G_{xy}(z) \sim C_{xy}(R - z)^{1/2} \qquad \text{if } R\mu(R) < 1.$$

PROOF. By Corollary 5.2, all of the Green's functions $G_{xy}(z)$ are algebraic; and by Proposition 7.4, all have the same singularity type at their common radius R of convergence. Consequently, all have convergent Puiseux expansions at $z = R$. It must be shown that the leading terms of these expansions are as implied by (7.37)-(7.38).

Since all of the leading terms have the same exponent, by Proposition 7.4, it will suffice to consider a convenient positive linear combination of the entries of $\mathbf{G}(z)$. Thus, consider the function $F(z) = \alpha(z)^T\mathbf{G}(z)\mathbf{1}$, where $\alpha(z)^T$ is the left Perron-Frobenius eigenvector of $M(z)$ and $\mathbf{1}$ is the vector of all 1s. By equation (7.20),

$$F(z) = (1 - z\mu(z))^{-1}.$$

Hence, by Corollary 7.18,

$$F(z) \sim C_F(R - z)^{-1/2} \qquad \text{if } R\mu(R) = 1; \text{ and}$$

$$F(R) - F(z) \sim C_F(R - z)^{1/2} \qquad \text{if } R\mu(R) < 1.$$

Since $\lim_{z \to R-} \alpha(z) = \alpha(R)$, it follows that

$$\alpha(R)^T\mathbf{G}(z)\mathbf{1} \sim C_F(R - z)^{-1/2} \qquad \text{if } R\mu(R) = 1; \text{ and}$$

$$\alpha(R)^T\mathbf{G}(R)\mathbf{1} - \alpha(R)^T\mathbf{G}(z)\mathbf{1} \sim C_F(R - z)^{1/2} \qquad \text{if } R\mu(R) < 1.$$

The relations (7.38)-(7.37) clearly follow. $\square$

COLLARY 7.20. *For any two words x, y of lengths $|x|, |y| \leq 1$, there exists a constant $C_{xy} > 0$ such that, as $n \to \infty$,*

(7.40) $P^x\{X_n = y\} \sim C_{xy} R^{-n}$ *if $R < \rho$;*

(7.41) $P^x\{X_n = y\} \sim C_{xy} R^{-n} n^{-1/2}$ *if $R = \rho$ and $R\mu(R) = 1$; and*

(7.42) $P^x\{X_n = y\} \sim C_{xy} R^{-n} n^{-3/2}$ *if $R = \rho$ and $R\mu(R)$ < 1.*

PROOF. The first statement has been proved in Corollary 7.10. In view of Corollary 5.2 and Proposition 7.19, the latter two results are immediate consequences of Corollary 4.1.   □

8. Transience and the spectral radius

By Corollary 7.20, the only possible limiting behaviors of the transition probabilities are (7.40), (7.41), and (7.42). Thus, to complete the proofs of Theorems 2.5-2.6, we must show that $R = 1$ if and only if the the RWRL is recurrent, because

(a) if $R = 1$, then (7.40) implies that the transition probabilities converge to nonzero limits, and so the RWRL is *positive* recurrent;
(b) if $R = 1$, then (7.41) implies that the transition probabilities converge to zero, and so the RWRL is *null* recurrent;
(c) if $R = 1$, then (7.42) implies that the transition probabilities are summable in n, and so the RWRL is *transient*, by a standard test for recurrence; and
(d) if the RWRL is recurrent, then the Green's functions are all infinite at $z = 1$, by the same test for recurrence, and so $R = 1$.

PROPOSITION 8.1. *If the RWRL is transient, then $R > 1$.*

It is possible to prove this in a purely "analytical" fashion, using Lemma 7.13 to show that $\rho = 1$ is impossible unless all the nondegenerate first-passage generating functions $H_{ab,c}(z)$ take the value 1 at $z = 1$, which implies that the RWRL is recurrent. However, it is simpler to use a probabilistic argument based on a standard large-deviations estimate, which we record in the following lemma.

LEMMA 8.2. *Let $\xi_1, \xi_2, \ldots$ be a sequence of Bernoulli random variables adapted to a filtration $\{\mathcal{F}_n\}_{n \geq 0}$. Assume that there exists $p > 0$ such that, for every $n \geq 1$,*

(8.1) $P(\xi_{n+1} = 1 \,\|\, \mathcal{F}_n) \geq p.$

Then for every $\alpha < p$ there exist $\beta < 1$ and $C < \infty$ such that $\forall m \geq 1$,

(8.2) $P\{\sum_{j=1}^{n} \xi_j \leq \alpha n \text{ for some } n \geq m\} \leq C\beta^m.$

PROOF. See [**8**].   □

Proof of Proposition 8.1. If the RWRL X_n is transient, then for every $k \geq 1$, with probability one, the first k letters of X_n eventually stabilize. Consequently, there exists a word $w \in \mathcal{L}$ such that

(8.3) $P^w\{X_n \text{ is prefixed by } w \,\forall n \geq 0\} := p > 0.$

Define $\mathcal{L}(w)$ to be the set of all words $v \in \mathcal{L}$ that are *suffixed* by w, i.e., $\mathcal{L}(w)$ is the set of all words in $\mathcal{L}$ of the form uw. By Assumption 2.4, for any word $w' \in \mathcal{L}$, there exists a positive-probability path γ from w' to a word $v \in \mathcal{L}(w)$ through

words of length $\geq |w'|$. Because only the last two letters in any transition influence the transition probabilities, the probability of such a path γ depends on w' only through its last two letters. Since there are only finitely many two-letter words in $\mathcal{L}$, it follows that, for some $\alpha > 0$ and some positive integer N,

$$(8.4) \qquad P^v\{X_n \in \mathcal{L}(w) \text{ for some } n \leq N\} \geq \alpha \ \forall v \in \mathcal{L}.$$

Consider the event $\{X_n = \emptyset\}$. In order that this event occur, it must be the case that, for every visit to the set $\mathcal{L}(w)$ prior to time n, the last letter of the suffix w is erased, or one of the last two letters modified, before time n. By (8.4) and Lemma 8.2, the probability that there are fewer than $\alpha n/2N$ visits to $\mathcal{L}(w)$ by time n decays exponentially in n. By (8.3) and Lemma 8.2, the probability that one of the the last two letter of the suffix w is modified $\alpha n/2N$ times before time n decays exponentially in n. Therefore, $P^\emptyset\{X_n = \emptyset\}$ decays exponentially in n, equivalently, $R > 1$.

Concluding Remark. While this paper was being reviewed, the author received a preprint of a paper by T. NAGNIBEDA and W. WOESS in which results similar to Theorems 1 and 2 are proved. The results of Nagnibeda and Woess are proved only for random walks on trees, but by a similar method.

References

[1] K. Aomoto (1984) Spectral theory on a free group and algebraic curves. *J. Fac. Sci. Univ. Tokyo* IA **31**, 297–317.

[2] E. Bender (1974) Asymptotic methods in enumeration. *SIAM Review*, **16**, 485–513.

[3] P. Bougerol (1979) Thorme central limite local sur certains groupes de Lie. *Ann. Sci. cole Norm. Sup.* (4) **14** 403–432.

[4] P. Bougerol & T. Jeulin (1999) Brownian bridge on hyperbolic spaces and on homogeneous trees. *Probability Theory & Related Fields* **115** 95–120.

[5] T. H. Cormen, C. E. Leiserson, & R. L. Rivest (1990) *Algorithms.* MIT Press, Cambridge MA

[6] Cox, Little, & O'Shea (1992) *Ideals, Varieties, and Algorithms.* Springer, NY

[7] W. Feller (196) *An Introduction to Probability Theory and its Applications.* Wiley, NY

[8] D. Freedman (1975) On tail probabilities for martingales. *Annals of Probability* **3** 100–118.

[9] A. Gairat, V. Malyshev, M. Menshikov, & K. Pelikh (1995) Classification of Markov chains that describe the evolution of random strings. (Russian) *Uspekhi Mat. Nauk* **50** (302)5–24; translation in *Russian Math. Surveys* **50** (1995) 237–255.

[10] P. Gerl & W. Woess (1986) Local limits and harmonic functions for nonisotropic random walks on free groups. *Probability Theory & Reated Fields*, **71**, 341–355.

[11] E. Hille (1959) *Analytic Functions*, volumes 1–2. Ginn, Boston.

[12] J. Hopcraft & J. Ullman (1979) *Introduction to Automata Theory, Languages, and Computation.* Addison-Wesley, Reading MA

[13] K. Kendig (1977) *Elementary Algebraic Geometry.* Springer-Verlag, NY

[14] S. Lalley (1993) Finite range random walk on free groups and homogeneous trees. *Annals of Probability* **21** 2087–2130.

[15] S. Lalley (1995) Return probabilities for random walk on a half-line. *J. Theoretical Probability* **8** 571–599.

[16] J. Lehner (1964) *Discontinuous Groups and Automorphic Forms.* American Math. Soc., Providence RI

[17] E. Le Page & M. Peigné (1999) Local limit theorems on some non-unimodular groups. *Rev. Mat. Iberoamericana* **15** 117–141.

[18] V. Malyshev (1994) Stabilization laws in the evolution of a random string. (Russian) *Problemy Peredachi Informatsii* **30** 79–95 translation in *Problems Inform. Transmission* **30** (1995) 260–274.

[19] M. Reed & B. Simon (1978) *Methods of Mathematical Physics IV: Analysis of Operators*. Academic Press, NY

[20] S. Sawyer (1978) Isotropic random walks on a tree. *Z. Warscheinlichkeitstheorie.*, **42**, 279–292.

[21] E. Seneta (1981) *Nonnegative Matrices and Markov Chains*, 2nd ed., Springer, NY.

[22] W. Woess (1982) A local limit theorem for random walks on certain discrete groups. *Probability measures on groups (Oberwolfach, 1981)*, pp. 467–477, Lecture Notes in Math. **928**, Springer, Berlin-New York.

[23] W. Woess (1987) Context-free languages and random walks on groups. *Discrete Math.* **67** 81–87.

UNIVERSITY OF CHICAGO, DEPARTMENT OF STATISTICS, USA
E-mail address: `lalley@galton.uchicago.edu`